

SEARCH REQUEST FORM

Requestor's Serial
 Name: _____ Number: _____
 Date: _____ Phone: _____ Art Unit: _____

Search Topic:

Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevant citations, authors keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevant claim(s).

STAFF USE ONLY

Date completed: ~~10~~ 11-04-02
 Searcher: Beverly E 4999
 Terminal time: 20
 Elapsed time: _____
 CPU time: _____
 Total time: 25
 Number of Searches: _____
 Number of Databases: 1

Search Site

_____ STIC
 _____ CM-1
 _____ Pre-S

Type of Search

_____ N.A. Sequence
 _____ A.A. Sequence
 _____ Structure
 _____ Bibliographic

Vendors

_____ IG Suite
 _____ STN
 _____ Dialog
 _____ APS
 _____ Geninfo
 _____ SDC
 _____ DARC/Questel
☒ Other CGN

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2002, 22:58:32 ; Search time 2426 Seconds
(without alignments)
10497.772 Million cell updates/sec

Title: US-09-720-384A-3
Perfect score: 1217
Sequence: 1 gcgcctttcattcatca.....aaaaaaaaaaaaaaaaaaaaa 1217

Scoring table: IDENTITY_NUC
Gap 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues
Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg.hum.*
- 31: em.htg_inv.*
- 32: em.htg_other.*
- 33: em.htg_inv.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description

1	311	25.6	970	8	AF178976	AF178976	Zea mays
2	309	25.4	1086	8	AF044285	AF044285	Catharant
3	296.4	24.4	1077	8	ATU05238	ATU05238	Arabidopsis
4	296.4	24.4	1143	8	AY054287	AY054287	Arabidopsis
5	296.4	24.4	1185	8	ATCLAPS	ATCLAPS	A.thaliana
6	263.6	21.7	1243	8	AF462823	AF462823	Arabidopsis
7	263.6	21.7	1311	8	AF043351	AF043351	Arabidopsis
8	170.2	14.0	1131	1	AE007523	AE007523	Clostridi
9	168.6	13.9	220050	1	AJ414156	AJ414156	Yersinia
10	163.4	13.4	51860	8	AB013390	AB013390	Arabidops
11	161.4	13.3	53533	1	BSY09476	BSY09476	B.subtilis
12	161.4	13.3	210440	1	BSUB0006	BSUB0006	Bacillus su
13	156.8	12.9	154330	2	AC104429	AC104429	Oryza sat
14	156.4	12.9	1148	8	AF017216	AF017216	Saccharom
15	155.6	12.8	2378	3	AB036852	AB036852	Ciona int
16	154.8	12.7	2080	8	AB049836	AB049836	Saccharom
17	153.8	12.6	1021	11	CNS06JUF	CNS06JUF	T7 end of
18	153.2	12.6	2014	9	AF091242	AF091242	Homo sapi
19	153.2	12.6	2377	9	AF150754	AF150754	Homo sapi
20	152.6	12.5	8135	1	BSPYREYLO	BSPYREYLO	Bacillus
21	152.6	12.5	208780	1	BSUB0009	BSUB0009	Bacillus su
22	152.2	12.5	2424	9	BC009894	BC009894	Homo sapi
23	152	12.5	1845	9	AF313907	AF313907	Homo sapi
24	152	12.5	1909	9	AF173365	AF173365	Homo sapi
25	152	12.5	3774	9	AF074331	AF074331	Homo sapi
26	151	12.4	3480	10	AF251798	AF251798	Cavia por
27	146	12.0	293181	1	AP001119	AP001119	Buchnera
28	143.2	11.8	1584	8	S55315	S55315	ATP-a
29	143.2	11.8	1584	8	SCMET14A	SCMET14A	S.cerevisia
30	143.2	11.8	2063	8	SCYKL001C	SCYKL001C	S.cerevisia
31	143.2	11.8	3876	8	SCMER17	SCMER17	S.cerevisia
32	143.2	11.8	24743	8	SCCEN1D	SCCEN1D	S.cerevisia
33	143	11.8	300550	1	AP001512	AP001512	Bacillus
34	141.6	11.6	303250	1	AP001518	AP001518	Bacillus
35	139.4	11.5	11117	1	AE003980	AE003980	Xylella f
36	139	11.4	990	23	E09951	E09951	cDNA encodi
37	139	11.4	1875	6	E27970	E27970	APS kinase/
38	139	11.4	1875	9	AF016496	AF016496	Homo sapi
39	139	11.4	2265	9	AF033026	AF033026	Homo sapi
40	139	11.4	2282	9	HSU53447	HSU53447	Homo sapien
41	139	11.4	2430	9	BC011392	BC011392	Homo sapi
42	139	11.4	2506	6	AR044160	AR044160	Sequence
43	139	11.4	2511	9	HSPAPSSYN	HSPAPSSYN	Sequence
44	139	11.4	2537	9	AF105227	AF105227	Homo sapi
45	139	11.4	2646	6	AX281701	AX281701	Sequence

ALIGNMENTS

RESULT 1	AF178976	AF178976	970 bp	mRNA	linear	PLN 13-DEC-1999
LOCUS	Zea mays	adenosine-5'-phosphosulfate kinase (AK1)	mRNA, partial			
DEFINITION	Zea mays	adenosine-5'-phosphosulfate kinase (AK1)	mRNA, partial			
ACCESSION	AF178976					
VERSION	AF178976.1	GI:6563284				
KEYWORDS						
SOURCE	Zea mays.					
ORGANISM	Zea mays					
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC					
AUTHORS	1 (bases 1 to 970)					
TITLE	Bolchi A., Petrucco, S. and Ottonello, S.					
JOURNAL	Isolation and comparative expression analysis of a maize cDNA encoding adenosine 5'-phosphosulfate kinase					
REFERENCE	Unpublished					
AUTHORS	2 (bases 1 to 970)					
TITLE	Bolchi A., Petrucco, S. and Ottonello, S.					
JOURNAL	Direct Submission					
TITLE	Submitted (19-AUG-1999) Istituto di Scienze Biochimiche,					
JOURNAL	Universita' di Parma, Viale delle Scienze, Parma 43100, Italy					

RESULT 2	AF044285				1086 bp	mRNA	linear	PLN 11-AUG-1998	
LOCUS	Catharanthus roseus adenosine-5'-phosphosulfate-kinase (Crakn)								
DEFINITION	mRNA, complete cds.								
ACCESSION	AF044285								
VERSION	AF044285.1				GI:2832299				
KEYWORDS	Madagascar periwinkle.								
SOURCE	Catharanthus roseus								
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Gentianales; Apocynaceae; Rauvolfioideae; Vincaceae; Catharanthus.								
REFERENCE	1 (bases 1 to 1086)								
AUTHORS	Arz,H.E., Gisselmann,G., Schiffmann,S. and Schwenn,J.D.								
TITLE	A cDNA for adenylyl sulphate (APS)-kinase from Arabidopsis thaliana								
JOURNAL	Biochim. Biophys. Acta 1218 (3), 447-452 (1994)								
MEDLINE	94325358								
REFERENCE	2 (bases 1 to 1086)								
AUTHORS	Schiffmann,S. and Schwenn,J.D.								
TITLE	Isolation of cDNA clones encoding								
	adenosine-5'-phosphosulfate-kinase (EC2.7.1.25) from Catharanthus								
	roseus (Accession No. AF044285) and an isoform (akn2) from								
	Arabidopsis (Accession No. AF043351) (PGR98-116)								
JOURNAL	Plant Physiol. 117 (3), 1125 (1998)								
AUTHORS	3 (bases 1 to 1086)								
TITLE	Schiffmann,S. and Schwenn,J.D.								
JOURNAL	Direct Submission								
	Submitted (23-JAN-1998) Biochemie der Pflanzen, Ruhr-Universitaet								
	Bochum, Universitaetsstrasse 150, Bochum 44780, Germany								
FEATURES	Location/Qualifiers								
source	1. .1086								
gene	/organism="Catharanthus roseus"								
	/db_xref="taxon:4058"								
	/sub_clones="KSCrakn"								
	/note="heterotrophic cell culture grown under sulfate								
	deficiency"								
	1. .1086								
	/gene="Crakn"								
CDS	39. .977								
	/gene="Crakn"								
	/EC_number="2.7.1.25"								
	/function="phosphorylates 3'-OH group of								
	adenosine-5'-phosphosulfate"								
	/note="APS-kinase; putative processing site after residue								
	46"								
	/codon_start=1								
	/product="adenosine-5'-phosphosulfate-kinase"								
	/protein_id="AAC31145.1"								
	/db_xref="GI:2832300"								
	/translation="MIGSVKRPVVCVLPPEFDFTSTGLKSSVKLPVNFQAFSGG								
	GGEVLGLFAPIKATEGSKTSFQVNGKVDNFRHLQPSDCNSDSLSNNGFPCKK								
	ILQTTVNGNTNLIWHKCAVEKSEKPEIQRCGVITITGLSGSKGLACALSLRLH								
	AKGLTYITLDGDNVVRHGLNSDLFSKAEADRAENIRRIGEVAKLFADAGVCIASLISPY								
	RKPPDACHSLPEGDFTEVPMQVPLKVCARDPKGLYKLARAGIKGFTGIDDPYEP								
	LKSEIVLHQKLGMCDSPLDADIVISYLENGYLKA"								
BASE COUNT	313 a	187 c	267 g	319 t					
ORIGIN									
Query Match	25.4%; Score 309; DB 8; Length 1086;								
Best Local Similarity	67.4%; Pred. No. 6.5e-63;								
Matches	435;	Conservative	0;	Mismatches	210;	Indels	0;	Gaps	0;
QY	398	CCGAGTCAAGGAGAACCCCTGTAATGTCCGAACATTTGGGAATCCACTAATATTTTATGCA	447						
DB	332	CCGAGTAAAGAGATATTACAAACCACTACCGTTGGAACTCTACTAATATACTTTGGCA	391						
QY	448	CAATTGCTTGAATGGACAATCTGTATAGACAGAAATTCCTGGGACAAAAGGCTGTGCTGT	507						
DB	392	TAAATGTGCTGTGAAAAAAGTGAAGCAGCAACCCCTGCACAGAAAGGCTGTGTTAT	451						

QY	508	ATGGATATACAGAGCTCAGTGGTTTCAGGAAAAAGTACTCTTCATGTCGACTGAGTCGTGA	567	CDS	33. .863	/BC_number="2.7.1.25"
Db	452	ATGGATTACTGCTTGTGGTTTCAGAAAGACAGCACTTGGCATGCTGTTTGGTCAGG	511			/function="phosphorylates 5'-adenylylsulfate on 3'-hydroxyl group"
QY	568	GTTCATTCAGAGCCACCTCAGCTATGCTTCTGATGTCGACCACTCAGACATGGCT	627			/note="similar to yeast APS kinase (MET14): SwissProt Accession Number Q02196; similar to Escherichia coli APS kinase (cysC): SwissProt Accession Number P23846; adenosine 5'-phosphosulfate kinase"
Db	512	CTTACATGCAAGGAAACCTAACCTACATCTTCTGATGCTGATAATGTTTCGACATGATT	571			/codon_start=1
QY	628	AAATAGAGATTTAAGCTTTAAGGCAGAGACCGTCGAGAAATATACGAAGATTCGGTGA	687			/evidence=experimental
Db	572	GAACAGTGATCTTAGTCTTAAAGCAGAGAAATCGAGCGAGAAATATTAGGAGGATTGGAGA	631			/product="APS kinase"
QY	688	AGTGGCAAGCTTTTGTGATGCTGGTGTCAATGTCATGCTAGCTTGATATCTCCATA	747			/protein_id="AAC50035.1"
Db	632	GGTGTCAAAGCTCTTGTGATGCTGGAGTCATTTTGCAATGTCGAAGTTTGATATCTCCCTA	691			/db_xref="GI:450235"
QY	748	CAGGAGAGATCGTATGATCGCCGCTCTACTTCCACATCTCACTTTATTTGAAGTATT	807			/translation="MIAAGKSLGLSMASPKGIFDSNMSNSRSVVVYACVSMDSQSLINKNGSIPVKSSINGHTGOKGPLSTGVNSTNKNWHECSVEKVDRLDDKGC
Db	692	CAGAAGCCCCGATGCTCTGCTCTTTGCTTCCAGAAAGGGATTTCATTGAGGTATT	751			VWVLTGSLGSKSLACALNQMLYQKGLCYLLDDGNVRRHGLNDRDLRSFRAEDRAENIR
QY	808	TATTGATTTGCCCTAAAAATTTGTGAAGCTCGTGATCCTTAAAGGCCTATACAAAGCTTGC	867			RYGEVAKLFADAGIICISLIISPYRDACRSLLPEGDFVEFMDVPLSVCEARDPK
Db	752	CATGGATGCTCCTCTGAAAGTCTCGGAGCAAGGACCTAAGGGCTTATATAAGCTTGC	811			GLYKLARAGKIKGFTGIDDPYEPPLNCEISLREGGTSPIEMAEKVYVYLDNKGYLQA
QY	868	ACGTACAGAAAGATTAAGGTTTCACTGGAATTTGATGATCCATACGACCAACCATTTAA	927			
Db	812	ACGAGCTGGTAAATCAAGAGCTTTACTGGAATTTGATGATCCTTACGAGCGGCCACTTAA	871			
QY	928	TGTTGATAGTAAATTAAGATCAAGATGAGGAATGCCCTTACCCCAAGCAATGGCCAA	987			
Db	872	ATCAGAGATGTTTTCATCAGAAAGCTTGGTATGTGTGATTTCTCCATGATTTGGCTGA	931			
QY	988	GCAAGTCTATGCTACCTTGAAGAAACGGATATTGCAAGCTTA	1032			
Db	932	TATAGTATATCATCTCGAGGAAATGATATCTAAAGGCATA	976			
RESULT 3						
ATU05238						
LOCUS		1077 bp	mRNA	linear	PLN 23-JUN-1998	
DEFINITION		Arabidopsis thaliana APS kinase mRNA, complete cds.				
ACCESSION		U05238				
VERSION		U05238.1	GI:450234			
KEYWORDS		thale cress.				
SOURCE		Arabidopsis thaliana				
ORGANISM		Arabidopsis thaliana				
REFERENCE		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.				
AUTHORS		Jain, A. and Leustek, T.				
TITLE		A cDNA clone for 5'-adenylylphosphosulfate kinase from Arabidopsis thaliana				
JOURNAL		Plant. Physiol. 105, 771-772 (1994)				
MEDLINE		94345022				
REFERENCE		2 (bases 1 to 1077)				
AUTHORS		Lee, S. and Leustek, T.				
TITLE		APS kinase from Arabidopsis thaliana: genomic organization, expression, and kinetic analysis of the recombinant enzyme				
JOURNAL		Biochem. Biophys. Res. Commun. 247 (1), 171-175 (1998)				
MEDLINE		98300303				
REFERENCE		3 (bases 1 to 1077)				
AUTHORS		Leustek, T.				
TITLE		Direct submission				
JOURNAL		Submitted (18-JAN-1994) Thomas Leustek, Center for Agricultural Molec. Biology, Rutgers University, Cook College, College Farm Road, New Brunswick, NJ 08903, USA				
FEATURES		Location/Qualifiers				
source		1..1077				
		/organism="Arabidopsis thaliana"				
		/db_xref="taxon:3702"				
		/clone="ATK"				
		/tissue_type="whole plant"				

transit_peptide 33. .143			
/note="potential chloroplast transit peptide with a protease cleavage site"			
mat_peptide		144..860	
misc_feature		354..377	
		/note="ATP/GTP binding site motif"	
BASE COUNT	300 a	178 c	256 g 343 t
ORIGIN			
Query Match	24.4%	Score 296.4;	DB 8; Length 1077;
Best Local Similarity	68.0%;	Pred. No. 6.5e-60;	
Matches	430;	Conservative	0; Mismatches 196; Indels 6; Gaps 1;
QY	411	TGTCGAACATTTGGAAATCGCACTAATATTTATGGCACAAATTCCTTGATGGACAATCTG	470
Db	247	TGTCACGGTCGGAAACTCGACAAATATAAAGTGGCATGAATGCTGTTGAGAAGTTG	306
QY	471	ATAGACAGAAATTCCTGGGACAAAAGCTGTGTCGTATGGATAACAGCACTCACTGCTT	530
Db	307	ATAGACAGACATTCCTTGATCATCAGAAAGGATGTGTGATTTGGGTCACCGCTCTTAGT	366
QY	531	CAGGAAAAGTACTCTTGATGTCGCTGAGTGTGAGTTGCGATTCGACAGGCCACCTCA	590
Db	367	CAGGAAAGAGTACTTTGGCTTGTGCTTGAATCAGATGTTGTATCAAAAGGGGAAGCTTT	426
QY	591	CGTATGCTGATGTCGACACCTCAGACATGGCCCTAAATAGAGATTTAAAGCTTTAAGG	650
Db	427	GTATATTTCTTGATGCTGATAATGTTAGGCATGGCTTAAACCGTGATCTTAGCTTTAAAG	486
QY	651	CAGAGACCGTCGAGAAAATATACGAAGAGTTGGTGAAGTGGCAAGCTTTTTCGTGATG	710
Db	487	CTGAGGATCGTCAGAGATATTCGTAGATTGGAGAGTTGCTAAGCTTTTTCGGGATG	546
QY	711	CTGGTGCATATGATGCTAGCTTGATATCTCCATACAGAGAGATCGTATGATGCC	770
Db	547	CTGGAATAATCTCATTTGCGAGTTTGATATCTCTTATAGAACAGATAGGGACGCTTGC	606
QY	771	GTGCTCTACTTCCACATTTCTAACTTTATTGAAGTATTATTATTATTTGCCCTTAAAAATTT	830
Db	607	GAAGTTTGTCTCCCGAGGAGAGATTTTGTGAGGTGTTTCATGGATGTACCCGCTTAGTGTTT	666
QY	831	GTGAAGCTCGTGATCTTAAAGCCCTATACAGCTTCACCTCAGCAAGAGATTAAGGTT	890
Db	667	CGSAGGCGAGGGATCCAAAGGGTCTTTTACAAGCTTGTCTGTCGAGGAAGATCAAGGTTT	726
QY	891	TCAGTGAATTTGATGATCCATACGAACCACTTAATATGTCGAGATAGTAATTAAGATGA	950
Db	727	TTACCGGATCGATGATACCCTTACGAGCCCACTTGAACCTCGAG-----ATTCTCTAG	780
QY	951	AAGATGAGGAATGCCCTTCACCAAGCAATGGCCAAAGCAAGTTCTATGCTACCTTTGAAG	1010
Db	781	GACGTGAAGGAGGAAGCTTCTCTCTATCGAAATGGCGGAAAAGGTCGTCGGATACCTTAGATA	840

Qy 1011 AAAACGATATTTGCAAGCTTACTATATATAT 1042
| | | | | | | | | | | | | | | | | | | |
Db 841 ACAAGGTTTATCTTCAAGCATACATCTCT 872

RESULT 4

AY054287

LOCUS

DEFINITION

ACCESSION

VERSION

FLI CDNA.

KEYWORDS

SOURCE

ORGANISM

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi

1 (bases 1 to 1143)

Chen, R., Chen, H., Kim, C.J., Koesema, E., Meyers, M.C., Ban, H.,

Bowser, L., Carninci, P., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y.,

Ishida, J., Jiang, P.X., Jones, T., Kamiya, A., Karlin-Neumann, G.,

Kawai, J., Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M.,

Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Pham, P.K.,

Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C.,

Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K.,

Davis, R.W., Theologis, A. and Ecker, J.R.

Arabidopsis cDNA clones

Unpublished

2 (bases 1 to 1143)

Chen, R., Chen, H., Kim, C.J., Koesema, E., Meyers, M.C., Ban, H.,

Bowser, L., Carninci, P., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y.,

Ishida, J., Jiang, P.X., Jones, T., Kamiya, A., Karlin-Neumann, G.,

Kawai, J., Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M.,

Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Pham, P.K.,

Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C.,

Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K.,

Davis, R.W., Theologis, A. and Ecker, J.R.

Direct Submission

Submitted (22-AUG-2001) Salk Institute Genomic Analysis Laboratory

(SIGAL), Plant Biology Laboratory, The Salk Institute for

Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,

USA

RIKEN Genomic Sciences Center (GSC) members carried out the

collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN

Arabidopsis Full-length cDNA') : Seki, M., Narusaka, M., Ishida, J.,

Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,

Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGECC (SSP) Consortium members carried out the

sequencing and annotation of the RAFL cDNAs: Chen, R., Chen, H.,

Kim, C.J., Koesema, E., Meyers, M.C., Shinn, P., Ban, H., Bowser, L.,

Dale, J.M., Goldsmith, A.D., Jiang, P.X., Jones, T., Karlin-Neumann, G.,

Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Nguyen, M.,

Onodera, C.S., Palm, C.J., Pham, P.K., Quach, H.L., Southwick, A.,

Tang, C.C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S.,

Davis, R.W., Theologis, A., and Ecker, J.R.

Chen, R. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to

this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk)

contributed equally to this work as PIs.

Location/Qualifiers

1. 1143

AY054287 1143 bp mRNA linear PLN 30-SEP-2001
Arabidopsis thaliana At2g14750/F26C24.11 mRNA, complete cds.
AY054287
AY054287.1 GI:15810037
FLI CDNA.
thale cress.
Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi

1 (bases 1 to 1143)

Chen, R., Chen, H., Kim, C.J., Koesema, E., Meyers, M.C., Ban, H.,

Bowser, L., Carninci, P., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y.,

Ishida, J., Jiang, P.X., Jones, T., Kamiya, A., Karlin-Neumann, G.,

Kawai, J., Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M.,

Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Pham, P.K.,

Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C.,

Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K.,

Davis, R.W., Theologis, A. and Ecker, J.R.

Arabidopsis cDNA clones

Unpublished

2 (bases 1 to 1143)

Chen, R., Chen, H., Kim, C.J., Koesema, E., Meyers, M.C., Ban, H.,

Bowser, L., Carninci, P., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y.,

Ishida, J., Jiang, P.X., Jones, T., Kamiya, A., Karlin-Neumann, G.,

Kawai, J., Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M.,

Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Pham, P.K.,

Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C.,

Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K.,

Davis, R.W., Theologis, A. and Ecker, J.R.

Direct Submission

Submitted (22-AUG-2001) Salk Institute Genomic Analysis Laboratory

(SIGAL), Plant Biology Laboratory, The Salk Institute for

Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,

USA

RIKEN Genomic Sciences Center (GSC) members carried out the

collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN

Arabidopsis Full-length cDNA') : Seki, M., Narusaka, M., Ishida, J.,

Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,

Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGECC (SSP) Consortium members carried out the

sequencing and annotation of the RAFL cDNAs: Chen, R., Chen, H.,

Kim, C.J., Koesema, E., Meyers, M.C., Shinn, P., Ban, H., Bowser, L.,

Dale, J.M., Goldsmith, A.D., Jiang, P.X., Jones, T., Karlin-Neumann, G.,

Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Nguyen, M.,

Onodera, C.S., Palm, C.J., Pham, P.K., Quach, H.L., Southwick, A.,

Tang, C.C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S.,

Davis, R.W., Theologis, A., and Ecker, J.R.

Chen, R. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to

this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk)

contributed equally to this work as PIs.

Location/Qualifiers

1. 1143

/organism="Arabidopsis thaliana"

/db_xref="taxon:3702"

/chromosome="2"

/clone="RAFL09-19-L13(R13881)"

/note="ecotype: Columbia"

1. 92

93. 923

/note="putative adenosine phosphosulfate kinase"

/codon_start=1

/product="At2g14750/F26C24.11"

/protein_id="AAL06946.1"

/db_xref="GI:15810038"
/translation="MIAAGAKSLLLGSMASPKGIFDSNSMSNRVVVVRACVSMDS
QTLSHNKGSIPEVKSINHTGKQGLPLSTVGNSTNKNHHCSEKVRDLRDLDDQGC
VWVTGLSSGSKSTLACALNQLYOKGLCYILDGDNVHGLNRLDSFKAEDRAENIR
RYGEVAKLIFADAGIICIASLISPYRTDRDACHSLLPEDGFVEFMDVPLSVCEARDPK
GLYKLARACKIKFTGIDDDPYEPPLNCELSLREGCTSPTEMAEKVYVGLDNKGYLQA
"

3'UTR 924. 1143
BASE COUNT 322 a 189 c 262 g 370 t
ORIGIN

Query Match 24.4%; Score 296.4; DB 8; Length 1143;
Best Local Similarity 68.0%; Pred. No. 6.6e-60;
Matches 430; Conservative 0; Mismatches 196; Indels 6; Gaps 1;

Qy 411 TGTGCAACATTTGGGAATCGACTAATATTTATGGCACAAATTTGATGGCAAACTG 470
| | | | | | | | | | | | | | | | | | | |
Db 307 TGTCTACGGTTCGGAACTCGACAAATATAAAGTGGCATGAATTTCTGTGAGAAAGTTG 366

Qy 471 ATACACAGAAATTTGCTGGGACAAAAGGCTGTCGTATGGATAACAGGACTCAGTGGTT 530
| | | | | | | | | | | | | | | | | | | |
Db 367 ATACACAGAAATTTGCTGGGACAAAAGGCTGTCGTATGGATAACAGGACTCAGTGGTT 426

Qy 531 CAGGAAAGAGTACTCTTTCATGTCACATGAGTTCGTGAGTTGTCATTCGAGAGCCACCTCA 590
| | | | | | | | | | | | | | | | | | | |
Db 427 CAGGAAAGAGTACTCTTTCATGTCACATGAGTTCGTGAGTTGTCATTCGAGAGCCACCTCA 486

Qy 591 CGTATGTACTTGTGATGTCACACCTTCAGCATGCGCTTAATAGAGATTTAAGCTTTAAGG 650
| | | | | | | | | | | | | | | | | | | |
Db 487 GTTATATCTTGTGATGTCACACCTTCAGCATGCGCTTAATAGAGATTTAAGCTTTAAGG 546

Qy 651 CAGGAAAGTCTGAGAAATATACGAGAGTTGGTGAAGTGGCAAAAGCTTTTGGCTGATG 710
| | | | | | | | | | | | | | | | | | | |
Db 547 CTGAGGATCTGAGAGAAATATTCGAGAGTTGGTGAAGTGGCAAAAGCTTTTGGCTGATG 606

Qy 711 CTGCTGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTC 770
| | | | | | | | | | | | | | | | | | | |
Db 607 CTGGAATACTGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTC 666

Qy 771 GTGCTCTACTTCCACATTTCACTTTATTTAGTATTTATTTAGTATTTATTTAGTATTT 830
| | | | | | | | | | | | | | | | | | | |
Db 667 GAAGTTTGTCTCCCGAGGAGAGATTTTGTGAGTGTTCATGAGTGTTCACGCTTGTGTT 726

Qy 831 GTGAAGTCTGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCAT 890
| | | | | | | | | | | | | | | | | | | |
Db 727 GCGAGGCGAGGATCCAAAGGCTCTTTACAGCTTGTCTGTCGAGAGAGATCAAGGTT 786

Qy 891 TCAGTGAATTTGATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATG 950
| | | | | | | | | | | | | | | | | | | |
Db 787 TTACCGGGATCGATGACCTTTACGAGCCACCATTTGAACCTGCGAG-----ATTTCTCTAG 840

Qy 951 AAGATGAGGAATGCCCTTCCACCAAGCAATGGCCAAAGTTCATGCTACCTTGAAG 1010
| | | | | | | | | | | | | | | | | | | |
Db 841 GACGTGAAGGAGAACTTCTCTCTATCGAAATGGCGGAAAGGTCGTCGAGTACTTAGATA 900

Qy 1011 AAAACGGATTTGCAAGCTTACTATATATAT 1042
| | | | | | | | | | | | | | | | | | | |
Db 901 ACAAGGTTATCTTCAAGCATACATATCTTCT 932

RESULT 5
ATCLAPSK

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi

1 (bases 1 to 1185)

Chen, R., Chen, H., Kim, C.J., Koesema, E., Meyers, M.C., Ban, H.,

Bowser, L., Carninci, P., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y.,

Ishida, J., Jiang, P.X., Jones, T., Kamiya, A., Karlin-Neumann, G.,

Kawai, J., Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M.,

Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Pham, P.K.,

Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C.,

Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K.,

Davis, R.W., Theologis, A. and Ecker, J.R.

Arabidopsis cDNA clones

Unpublished

2 (bases 1 to 1185)

Chen, R., Chen, H., Kim, C.J., Koesema, E., Meyers, M.C., Ban, H.,

Bowser, L., Carninci, P., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y.,

Ishida, J., Jiang, P.X., Jones, T., Kamiya, A., Karlin-Neumann, G.,

Kawai, J., Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M.,

Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Pham, P.K.,

Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C.,

Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K.,

Davis, R.W., Theologis, A. and Ecker, J.R.

Direct Submission

Submitted (22-AUG-2001) Salk Institute Genomic Analysis Laboratory

(SIGAL), Plant Biology Laboratory, The Salk Institute for

Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,

USA

RIKEN Genomic Sciences Center (GSC) members carried out the

collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN

Arabidopsis Full-length cDNA') : Seki, M., Narusaka, M., Ishida, J.,

Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,

Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGECC (SSP) Consortium members carried out the

sequencing and annotation of the RAFL cDNAs: Chen, R., Chen, H.,

AUTHORS Arz,H.E., Gisselmann,G., Schiffmann,S. and Schwenn,J.D.
TITLE A cDNA for adenylyl sulphate (APS)-kinase from Arabidopsis thaliana
JOURNAL Biochim. Biophys. Acta 1218 (3), 447-452 (1994)
MEDLINE 94325358
REMARK (sites)
REFERENCE 2 (bases 1 to 1185)
AUTHORS Arz,H.E., Gisselmann,G., Schiffmann,S. and Schwenn,J.D.
TITLE A chloroplast APS-kinase cDNA from Arabidopsis thaliana
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 1185)
AUTHORS Schwenn,J.D.
TITLE Direct Submission
JOURNAL Submitted (03-NOV-1993) Schwenn J. D., Ruhr University Bochum, Biology, Universitaetsstr. 150, 44780 Bochum, Germany
FEATURES
source
Location/Qualifiers
1..1185
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/clone="PM239x14cDNA"
/tissue_type="leaf"
/clone_lib="cDNA in lambda ZapII"
/dev_stage="vegetative"
42..872
/standard_name="Atp:adenylylsulfate-3'-phosphotransferase"
/EC_number="2.7.1.25"
/function="phosphorylates 3'-OH group of adenylylsulfate"
/citation={2}
/codon_start=1
/product="APS kinase"
/protein_id="CAA53426.1"
/db_xref="GI:414737"
/db_xref="SPTREMBL:Q43295"
/translation="MTAAKSLGLGSMASPKGIEDSNMSNSRVVVVACVSMDS
QTLVSHKNSISPEVKSINGHFCQKQGLSTVNSINIKWHCSVEKVDQRLLDQKGC
VITWGLSGSKSTLACALNOMLFQKGLCYILDGNVHRHLSRDLSPKAEADRAENIR
RVGEAKLADAGIICIASLSPYTRDRACSLPEGDFVEFMDPLSVYCARDPK
GLYKLARAGKIRGFTGIDDPYEPPLNCEISLREGGTSPIEMAEKVVGYLDNKGYLQA
"
transit_peptide 42..155
/note="putative"
mat_peptide 162..869
/EC_number="2.7.1.25"
/note="putative mature peptide only - gene product starting with Arg at position 162 is enzymatically active, but maturation in the chloroplast has not yet been shown. Homology to sequences from E. coli & ye"
/product="APS-kinase"
stem_loop 984..1012
polya_site 985..990
stem_loop 1092..1120
BASE COUNT 335 a 200 c 267 g 383 t
ORIGIN
Query Match 24.4%; Score 296.4; DB 8; Length 1185;
Best Local Similarity 68.0%; Pred. No. 6.6e-60;
Matches 430; Conservative 0; Mismatches 196; Indels 6; Gaps 1;
QY 411 TGTCCGAACATGGGAATCGCAATATATTTATGGCACAATGCTTCATGGACAATCTG 470
DB 256 TGCTACGGTCGGAATCGCAATATATAAGTGGCATGAATGTCTGTTGAGAAAGTTG 315
QY 471 ATAGACAGAAATTCCTGGGCAAAAAGCGTGTCTGTATGATAACAGGACTCAGTGGTT 530
DB 316 ATAGACAGAGATTCGTTGATCAGAAAGGATGTGATTTGGGTACC GGCTCTTAGTGGTT 375
QY 531 CAGGGAAGACTCTTGCATGTGCACCTGAGTCGTGAGTTCGATCGCAGAGGCCACTCA 590
DB 376 CAGGGAAGACTCTTGGCTTGTGCTTGAATCAGATGTTATCAAAAGGGAGAGCTTT 435
QY 591 CGTATGTACTTCTGGTGCACACCTCAGACATGGCCTTAATPAGAGATTTAAAGCTTAAAG 650
DB 436 GTTATATCTTGTGGTGATAATGTTAGGCATGGCTTAAACCGTGATCTAGCTTTAAAG 495

QY 651 CAGAAGACCGTCAGAAAAATATACAAAGAGTTGGTGAAGTGGCAAAGCTTTTCTGTGATG 710
DB 496 CTGAGGATCGTCAGAGAATATTCGTAGAGTTGGACAGAGTTCTGCTAAGCTTTTTCGGGATG 555
QY 711 CTGCTGTATATGTCATTCGTAGCTTGTATCTCCATACAGAGAGATCGTCGATCGATGCC 770
DB 556 CTGGAATATCTGCATTCGGAGTTTGATATCTCTTATAGAACAGATAGGAGCGCTTGTG 615
QY 771 GTGCTCTACTTCCACATTTCTAACTTTTATGAAGTATTTATGATTTGCCCTCAAAAATTT 830
DB 616 GAAAGTTTGTCTCCCGAGGGAGATTTTCTTGAGGTGTTCATGGATGTACCGCTTAGTGTCTT 675
QY 831 GTGAAGCTGTGATCCTTAAAGCGCTATACAAAGCTTCACGTACAGAAAGATTTAAAGGTT 890
DB 676 GCAGCGGAGGATCCAAAGGCTCTTTACAAGCTTCTGCTGCAGAAAGATCAAGGTT 735
QY 891 TCACTGGAATGATGATCCATACGACCAACCAATTAATGGTGAGATAGTAATTAAGATGA 950
DB 736 TTACCGGGATCGATGACCTTTACGAGCCACCATTTGAATCGCGAG-----ATTCTCTAG 789
QY 951 AAGATGAGGAATGCCCTTACCCAAAGCAATGGCCAAAGCTTCTATGCTACCTTTGAAG 1010
DB 790 GACGTGAAGSAGGAATCTCTCTATCGAATGGCGAAAGGTCGTCGATACTTAGATA 849
QY 1011 AAAACGGATATTGGCAAGCTTAGTATATGTAT 1042
DB 850 ACAAGGTTATCTTCAAGCATAAACATATTCT 881
RESULT 6
AF462823 1243 bp mRNA linear PLN 08-JAN-2002
LOCUS Arabidopsis thaliana AF462823.1
DEFINITION Arabidopsis thaliana AF462823.1
ACCESSION AF462823.1
VERSION AF462823.1
KEYWORDS FLI_CDNA,
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi
s (bases 1 to 1243)
Cheuk,R., Chen,H., Kim,C.J., Meyers,M.C., Banh,J., Bowser,L.,
Carninci,P., Chang,E., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y.,
Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J.,
Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M.,
Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M.,
Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K.,
Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A.
and Ecker,J.R.
Arabidopsis cDNA clones
Unpublished
2 (bases 1 to 1243)
Cheuk,R., Chen,H., Kim,C.J., Meyers,M.C., Banh,J., Bowser,L.,
Carninci,P., Chang,E., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y.,
Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J.,
Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M.,
Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M.,
Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K.,
Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A.
and Ecker,J.R.
Direct Submission
Submitted (27-DEC-2001) Salk Institute Genomic Analysis Laboratory
(SIGNAL), Plant Biology Laboratory, The Salk Institute for
Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
USA
RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
Arabidopsis Full-Length cDNA') : Seki,M., Narusaka,M., Ishida,J.,
Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
Hayashizaki,Y. and Shinozaki,K.
The Salk, Stanford, PGEC (SSP) Consortium members carried out the

sequencing and annotation of the RAFL cDNAs: Cheuk, R., Chen, H., Kim, C. J., Meyers, M. C., Shinn, P., Banh, J., Bowser, L., Dale, J. M., Goldsmith, A. D., Jiang, P. X., Jones, T., Karlin-Neumann, G., Lam, B., Lee, J. M., Lin, J., Liu, S. X., Miranda, M., Nguyen, M., Onodera, C. S., Palm, C. J., Pham, P. K., Quach, H. L., Southwick, A., Tang, C. C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Davis, R. W., Theologis, A., and Ecker, J. R.

Cheuk, R. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Ecker, J. R. (SSP/Salk) contributed equally to this work as PIs.

FEATURES		Location/Qualifiers	
source	1..1243	/organism="Arabidopsis thaliana"	
		/db_xref="taxon:3702"	
		/chromosome="4"	
		/clone="RAFL09-09-101(R21470)"	
5'UTR	1..69	/note="ecotype: Columbia"	
	70..951		
		/note="adenosine-5'-phosphosulfate-kinase"	
		/codon_start=1	
CDS		/product="At4g39940/T5J17_110"	
		/protein_id="AAL58913.1"	
		/db_xref="GI:18087563"	
		/translation="MEGLAIRASRPSVFCISIPGLGDSHRKPPSDGFLKLPASSIPAD SRKLVANSTSFHPISAVNVSAQSLTADFPALSETILKEGRNNGKEAKNIIVHSESI CRCDROLLOQKGCYVWITGLSGSKSTVACALSKALPERGKLTITLDGDNVHGLNR DLTFAKHRTENIRIGEIVAKLPADVVICIASLISPYRRDRACRSLLPDGDFVEF MDVPLHVCESRDPKLYKLARAGKIKFTGIDDPYEAPVNCVVLKHTGDDSCSPRQ MAENIISYLNKGYLEG"	
3'UTR	952..1243		
	BASE COUNT	341 a	276 c 285 g 341 t
	ORIGIN		
	Query Match	21.7%;	Score 263.6; DB 8; Length 1243;
Matches	Best Local Similarity	65.5%;	Pred. No. 4.3e-52;
	402; Conservative	0;	Mismatches 209; Indels 3; Gaps 1;
	QY 422	GGGAATCGACTATATTTATGGCACAAATTCCTTGATTGGACAANTCTGATAGACAGAA	481
	Db 337	GAGAAGCAGAGAACATCGTGTGGCAGCAGAGTTCGATATGCGATGCGCAGACAGACAA	396
QY 482	TGTCGTGCACAAAGAGCTGTGCTGATGATTAACAGGACTCAAGTGTTCAGGGAAAGT	541	
Db 397	CTTCTTCAACAAAGAGGTTCTGCTGTTGGATCACTGCTCAAGTGTTCAGGGAAAGC	456	
QY 542	ACTCTTGACTGTGACGTGCTGATGTTGCATTGCAGAGGCCACCTCAGTATGACTT	601	
Db 457	ACTGTTGCTTGTGCATTAAGCATTTGTTGAAAGAGGCAAACTTACTTACACACTC	516	
QY 602	GATGTGCACACCTTCAGACATGGCCTTAATAGAGATTTAAGCTTTAAGGCAGAGACCGT	661	
Db 517	GACGCGACAAATGCTCGCTACCGCCTTAACCGTGACCTCACTTTCAAGCAGACGCCGC	576	
QY 662	GCAGAAATATACAGAGTTGGTGAAGTGCAGCAAGCTTTTTCGTGATGCTGGTGCATA	721	
Db 577	ACCGAAACATTAAGAAATTTGGTGAAGTGGCTGAAGTGTGCTGACGTCGAGTCAAT	636	
QY 722	TGCATTGCTAGCTTGATATCTCCATACAGAGAGATGCTGATGCATGCGGTGCTCACTT	781	
Db 637	TGTATAGCAAGTTTGATTCTCGTACCGAGAGACAGACGCGTGCCGGTCTTGTTA	696	
QY 782	CCACATTCTACTTTATTGAAGTATTTATTGATTTGCCCTTAAAAATTTGTGAAGCTGT	841	
Db 697	CTGACGCGGATTTGCGCGAGGTCCTTCATGGACGTTCTCTTATGTTGGCAGTCGAGA	756	
QY 842	GATCCTAAGCCCTATACAGCTTCGACGTACAGCAAGATTAAGGTTTCACTGGAATT	901	
Db 757	GATCAAGGGGTTGTACAGCTTGCAGCTGCAGGCAAAATCAAGGCTTCACTGGAATC	816	
QY 902	GATGATCCATACGACCAACCAATTAATGTTGAGATAGTATTA---AGATGAACATGAG	958	

Db 817	GACGACCCCTTACGAGGCACCAGTGAATTCGGAGGTAGTCTGCTCAAAACACAGACGAC	876
QY 959	GAATGCCCTTCACCCAAAGCAATGCCCAAGCAAGTTTCTATGCTACCTTGAAGAAACGGA	1018
Db 877	GATCGTGTTCGCCAGCTCAGATGGCTGAGAACATCATCTCTTACCTGCACAAACAAAGT	936
QY 1019	TATTTCCAAAGCTTA 1032	
Db 937	TATCTTGAGGGCTA 950	

RESULT 7	AF043351	1311 bp	mRNA	linear	PLN 11-AUG-1998
LOCUS	Arabidopsis thaliana adenosine-5'-phosphosulfate-kinase (akn2)				
DEFINITION	mRNA, complete cds.				
ACCESSION	AF043351				
VERSION	AF043351.1	GI:2829132			
KEYWORDS	thale cress.				
SOURCE	Arabidopsis thaliana				
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.				
REFERENCE	1 (bases 1 to 1311)				
AUTHORS	Arz, H. E., Gisselmann, G., Schiffmann, S. and Schwenn, J. D.				
TITLE	A cDNA for adenylyl sulphate (APS)-kinase from Arabidopsis thaliana				
JOURNAL	Biochim. Biophys. Acta 1218 (3), 447-452 (1994)				
MEDLINE	94325358				
REFERENCE	2 (bases 1 to 1311)				
AUTHORS	Schiffmann, S. and Schwenn, J. D.				
TITLE	Isolation of cDNA clones encoding adenosine-5'-phosphosulfate-kinase (EC2.7.1.25) from Catharanthus roseus (Accession No. AF044285) and an isoform (akn2) from Arabidopsis (Accession No. AF043351) (PGR98-1116)				
JOURNAL	Plant Physiol. 117 (3), 1125 (1998)				
REFERENCE	3 (bases 1 to 1311)				
AUTHORS	Schiffmann, S. and Schwenn, J. D.				
TITLE	Direct Submision				
JOURNAL	Submitted (18-JAN-1998) Biochemie der Pflanzen, Ruhr-Universitaet Bochum, Universitaetsstrasse 150, Bochum 44780, Germany				
FEATURES	Location/Qualifiers				
source	1..1311				
	/organism="Arabidopsis thaliana"				
	/cultivar="Columbia"				
	/db_xref="taxon:3702"				
	/clone="pGAD10"				
	/tissue_type="green vegetative"				
	/dev_stage="3 week old"				
gene	1..1311				
	/gene="akn2"				
CDS	61..942				
	/gene="akn2"				
	/EC_number="2.7.1.25"				
	/function="phosphorylates 3'-OH group of adenosine-5'-phosphosulfate"				
	/note="APS-kinase isoform; similar to the product encoded by GenBank Accession Number X75782"				
	/codon_start=1				
	/product="adenosine-5'-phosphosulfate-kinase"				
	/protein_id="AAC39520.1"				
	/db_xref="GI:2829133"				
	/translation="MEGLAIRASRPSVFCISIPGLGDSHRKPPSDGFLKLPASSIPAD SRKLVANSTSFHPISAVNVSAQSLTADFPALSETILKEGRNNGKEAKNIIVHSESI CRCDROLLOQKGCYVWITGLSGSKSTVACALSKALPERGKLTITLDGDNVHGLNR DLTFAKHRTENIRIGEIVAKLPADVVICIASLISPYRRDRACRSLLPDGDFVEF MDVPLHVCESRDPKLYKLARAGKIKFTGIDDPYEAPVNCVVLKHTGDDSCSPRQ MAENIISYLNKGYLEG"				
	238..240				
	/gene="akn2"				
	/note="putative processing site 3' of this position; unclassified site"				
misc_feature	373 a	293 c	303 g	341 t	1 others
BASE COUNT	373 a	293 c	303 g	341 t	

ORIGIN	2 (bases 1 to 11131)	REFERENCE
Query Match	21.7%; Score 263.6; DB 8; Length 1311;	AUTHORS
Best Local Similarity	65.5%; Pred. No. 4.4e-52;	TITLE
Matches 402; Conservative	0; Mismatches 209; Indels 3; Gaps 1;	JOURNAL
		FEATURES
QY 422 GGGAAATCGACTAATATTTATGCGACAATTCCTTGATTGGACATCTGATACAGAAA 481		source
Db 328 GAGAAAGCAGAGAACTCGTGGCAGAGAGTTTCATATGACATGCGACAGACAA 387		
QY 482 TTGCTGGGACAAAAAGCTGTGCTATGGATAACAGACACTCAGTGGTTCAGGAAAAAGT 541		gene
Db 388 CTTCCTTCAACAAAAGGGTGTGCTTGGATCACTGGTCTCAGTGGTTCAGGAAAAAGC 447		
QY 542 ACTCTTGTCATGTCACCTGAGCTGAGTTCATTCAGAGCGCCACCTCAGCTATGTACTT 601		CDS
Db 448 ACTGTTGCTGTGCACTAAGTAAAGCAATGTTTGAAGAGGCAACTTACTTACACTC 507		
QY 602 GATGGTGACAACTCAGACATGGCCTAAATAGAGATTTAAAGCTTTAAGGCAGAGACCGT 661		
Db 508 GACGGGGACAATGTCCTGTCAGCGCTTAACCGTGACCTCACTTTCAAGCAGAGCACGCG 567		
QY 662 GCAGAAATATACGAAGAGTTGGTGAAGTGGCAAGCTTTTGTGATGCTGGTGTCATA 721		
Db 568 ACCGAAAACATTAGAGAAATTTGGTGGCTAAGTGTGTTGCTGACGTCGGAGTCAT 627		
QY 722 TGCATGTGCTGATATCTCCATACAGAGAGATCGTGATGATGCCGTGCTTCACTT 781		gene
Db 628 TGTATAGCAAGTTTGAATTTCTCCGTACCGGAGACAGACAGCGCTGCCGCTCCTTGT 687		
QY 782 CCACATCTCAACTTTATTGAAGTATTATTGATTTGCCCTTAAATAATTTGTGAAGCTCGT 841		CDS
Db 688 CCGTACGGCGAATTCGTGAGGCTTCATGGACGTTCTCTTCATGTCGCGAGTCGAGA 747		
QY 842 GATCCTAAGGCGCTATACAGCTTGCACGTACAGGAAGATTAAGGTTTCACCTGGAATT 901		
Db 748 GATCCAAAGGGGTTGTACAGCTTGCACGTGCGAGCAAAATCAAGGCTTCACCTGGAATC 807		
QY 902 GATGATCCATAGCAACCACTAATTAATGTGTAGATAGTAATTA---AGATGAAGATGAG 958		gene
Db 808 GACGACCCCTACGAGGACCAAGTATGCGAGGTAGTCTCGAACACACAGACGACGAC 867		
QY 959 GAATGCCCTTACCACAAAGCAATGCCAGCAAGTCTTATGCTACCTTGAAGAAACGGA 1018		
Db 868 GAGTCGTGTTCGCCACGTGAGATGGCTGAGAACATCATCTCTTACCTGCCAAACAAAGGT 927		
QY 1019 TATTGCAAGCTTA 1032		
Db 928 TATCTTGAGGGCTA 941		
RESULT 8		
AE007523	11131 bp DNA linear BCT 27-JUL-2001	
LOCUS		
DEFINITION	Clostridium acetobutylicum ATCC824 section 11 of 356 of the complete genome.	
ACCESSION	AE007523	
VERSION	AE001437	
KEYWORDS	AE007523.1 GI:15022925	
SOURCE		
ORGANISM	Clostridium acetobutylicum. Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae; Clostridium.	
REFERENCE	1 (bases 1 to 11131)	
AUTHORS	Nolling,J., Breton,G., Omeichenko,M.V., Markarova,K.S., Zeng,Q., Gibson,R., Lee,H.M., Dubois,J., Oiu,D., Hitti,J., Wolf,Y.I., Tatusov,R.L., Sabathe,F., Doucette-Stamm,L., Soucaille,P., Daly,M.J., Bennett,G.N., Koonin,E.V. and Smith,D.R.	
TITLE	Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clostridium acetobutylicum	
JOURNAL	J. Bacteriol. 183 (16), 4823-4838 (2001)	
MEDLINE	21359325	
PUBMED	11466286	

```

GKVTAGFSGVKEAKVYIIYAKAVICATGGAAGIYKPNPNPFGSRHKMWYSPFNTGACY
AMGIRAGAEWTFEMRFTALRCKDITIAPTGCTIAAGICTPOVNAFGEKVTYRYGKPTTGI
NRLYSTIMENLNGPCVLYKTGTEESOVODLFKAVLNMPAQALKWEDRDINPSKON
IEITEPTPIYGHGASGYWDTTRKTTLDGLYAAGDVSGGSPKKYVYTGCPAEGEITAA
ENAEITSNKHLNITIDDEIDEITHRIKDFDGKSDFKIEQLEEARQWQYDETAGGIS
QHYRYTEKLNIAERIEBELLSYKLIKAKGSHELLFIFELIDRLYVAKVYVIAHLKAR
RTRWKCYNENESYTARDENFLKYVNSIYKEGKNVILYRLVGRNEIYEH"
5546. .5860
/gene="CAC0105"
CDS
5546. .5860
/gene="CAC0105"
/codon_start=1
/transl_table=11
/product="ferredoxin"
/protein_id="AAK78090.1"
/db_xref="GI:15022930"
/translation="MSIKIDLKVCVGQKCNICPGSLIDKNNDKAVIKYPKDCWGC
TACLKECKFKAIKFLGADIGGSGYMYVSEKGEELHWMINEDGKEIITTSRKESN
KY"
5802. .6908
/gene="CAC0106"
CDS
5802. .6908
/gene="CAC0106"
/codon_start=1
/transl_table=11
/product="ABC-type probable sulfate transporter,
periplasmic binding protein"
/protein_id="AAK78091.1"
/db_xref="GI:15022931"
/translation="MRMARKRLRLVGRNOININIKGSEVKKIKIGVLLSTIFTIG
ILSCGTESEKLSOVKIAYFGNITHSQALLORDGSLQKALGNSPKIKWEKFSAGSE
VESLLEGVDTIGTGFPAINAYTKNSGDIQYTAGADAGAILVSKKGADIKSVKDK
NKRVAIPQNTQDUTLRILLYENGLDRAGKNVEIASENSDIKTLIDGSGSIDAAL
YFPWFTIRVNEVAKVVLINIEDIWRKQYPTAIYVAFKEFYKAHPDIIVERFYKNLVE
ETKRVNSRDNSEKAINRQLKEITGKGLSKLIDSSFERIKVTNNPDERDAVLDMNWS
FKAGFTKSKADLKMENLSFLNRALKESGKAEIK"
6945. .7784
/gene="CAC0107"
CDS
6945. .7784
/gene="CAC0107"
/codon_start=1
/transl_table=11
/product="ABC-type sulfate transporter, AtPase component"
/protein_id="AAK78092.1"
/db_xref="GI:15022932"
/translation="MGIEIKGVSKDFISKEKKVHTLDEVNLSIKKGEFICLLGPSGCG
KSTLLNIAGLEKPTKGIYVLNNMEIKAPGPDRAFPQESALFPWLKVIDNVBEFKMI
KQVPKEREKALKYLVKWHLTKQNSYVHELSCGMQRFVALARALTLDSEVLLMDP
FSALDSQTSILOLELQNIWHTKTTIVFVTHNVEAVILLADRVVYVMAANPGRKNIF
EIKLARPRAQSVDLSYVAREVMRDLKEEVERKAKTEYDSDWNFQKSSVLNDSUDLG
ANL"
7696. .8493
/gene="CAC0108"
CDS
7696. .8493
/gene="CAC0108"
/codon_start=1
/transl_table=11
/product="ABC-type probable sulfate transporter, permease
protein"
/protein_id="AAK78093.1"
/db_xref="GI:15022933"
/translation="MQKRSTTVTGTFRRVFLYIILAIWELIYKVFDVLGVKMYFTF
PSPEVYNALAYLVKDNLTFIATCASLTRIVIGYIIALFIGIALALVCIKYKVDENL
TPTMLGLQTLPSVCNTPFALLWYCIDERSIIFVTYVGSLSFIAISTISGVKNVNIYV
KAARTWAGRGKNLYFDVYVIPSALPELISGMROQSFANRALMAGEMLSATKGLQVLT
AGRELGDTSQVWMMIIIIAFGLFDLFLFERVEKNIRYKCGLEKKN"
8506. .9306
/gene="CAC0109"
CDS
8506. .9306
/notes="cysd"
/gene="CAC0109"
/codon_start=1
/transl_table=11

```

```

/product="Sulfate adenylate transferase, CysD subfamily"
/protein_id="AAK78094.1"
/db_xref="GI:15022934"
/translation="MOHLOPLEAESITYILREAYHKFGKGLMLSGICKSDTVMLLAEK
AFFGNCPLPHVDVTHKIPEMIKRDEYAKYDLNLIVNTMDALNAGMGPHKGPLE
CCKALKUTDGOQVVDYTKFPELLIVGRDEDESGSKERSYFERNSESDWDYDQDPPEL
WQDFKTEFGKNHRIHPLHPLWELDIWIEYIRRENIIPVSLYFSKNGKRYRSLGSCAPC
TSPIDSNASTLDEIEELKNTNETSERSQAQDKADPHALEKLRKHGYM"
9320. .10888
/gene="CAC0110"
CDS
9320. .10888
/notes="cysn"
/gene="CAC0110"
/codon_start=1
/transl_table=11
/product="GTPase, sulfate adenylate transferase subunit 1"
/protein_id="AAK78095.1"
/db_xref="GI:15022935"
/translation="MSRENLNVFVGHVDHGKSTLIGRLLYDTNSLIPDGATEKVKKI
SAEKKGFYAFLLDAPFEBROQGITIDITMIQFFTKRDYVIDAPGHKEFLKNMIS
GAASAEAILVVDAGEIQOSKRGYILSLGKIKYVAVNKMOLVDVSEERYEIV
TQNSFLANLIVPEAYIPISAFGLDNVAKKSEKMPYKSKSLDMDSDVDEKGTEN
KALRFPQDIYKFDNRRIAGRIESGTLKEGDEIVFVPSGKTKVKASVEPDEKQKK
QVSARMSVGIITVEDEFYNNRGEIICHKNDGVTVSNFANFVWNGEPEKGGKTKLK
TATQIEAEAEVEIKKYIDAATLEEITGADHINKNDVAEIVFKSKKPCDFADNENAL
GRFVIDNYNTSGGIIILKLETYRDNKEISLVTKEEREKKFGSGKMKIVINGENKE
VTYQFALKLERKFLDQRAYFLASDESIAEILLRENGIIVIVANKEKSDSEYIEITY
DKDGKISDNILGII"
4368 a 1256 c 2302 g 3205 t
BASE COUNT
ORIGIN
Query Match 14.0%; Score 170.2; DB 1; Length 11131;
Best Local Similarity 58.4%; Pred. No. 1.2e-229;
Matches 317; Conservative 0; Mismatches 223; Indels 3; Gaps 1;
QY 415 GAACATTGGGAATGACACTAATATTTTATGGCACAAATTCCTTGATTGGACAATCTGATAG 474
DB 3266 GAATAATAAAAAAGATACAAATGTAGTGTGCAGGAACTATAAAATAAAAAGACAGATAG 3325
QY 475 ACAGAAATTCGCGGACAAAAAGCGTGTGTATGATGATACAGGACTCAGTGGTTACGG 534
DB 3326 AGAAAAATGTTTAAACAGAGGGTGCAGTTTTATGGTTTACTGGACTTTCAGGTTACGG 3385
QY 535 GAAAAGTACTCTTGGCATGTGCACTGAGTCGTGAGTTGCATTTGCAGAGGCCACTCACGTA 594
DB 3386 TAAGTCAACGGTTGTCATCTGCACTTGAAAGAAGCTTTATGAAATGGGTACTTCACATA 3445
QY 595 TGCTACTTGATGGTGCAACCTCAGACATGGCGCTAAATAGAGATTTTAAGCTTTTAAGGCAGA 654
DB 3446 TTTATTGGATGGAGACAAATTTAAGATATGGCTTTAATTCGTGATCTTGGCTTTAAAGTGA 3505
QY 655 AGACCGTGCAGAAAAATATACGAAGAGTTGGTGAAGTGGCGGCAAGCTTTTTCGTGATGCTGG 714
DB 3506 AGACAGAACTGAAAAATATAAGAAGAGTATCAGAAAGTTGCAAAAGCTGTTTTCGTGATGCG 3565
QY 715 TGTCAATATGCTTGTAGCTTCATATCTCCATACAGAGAGAGATGCTGATGCGCGTGC 774
DB 3566 AATCATTAACAAATAACAACATTTATTTCTCCATTCATAGAGGATAGAAATAATGCAAGAAA 3625
QY 775 TCTACTTCCACATCTTAACACTTTATTTTGAAGTATTTATTTGATTTTGGCCCTTAAAAATTTGTA 834
DB 3626 GCTTTTAAAGGAAA---AGATTTTGTAGAGGTATATATAGATTGTCCTTATTTGAAGTCTGTGA 3682
QY 835 AGCTCGGTGATCCTTAAAGCGCTATACAAGCTTGGACGTACAGGAAGAGATTTAAAGGTTTCCAC 894
DB 3683 GAAACGAGAGCCCCAAAAGGCATATATAAAAAAGCTAGAAAATGGTGAGATCAAAAATTTTAC 3742
QY 895 TGGAAATTTGATGTCATACGAACACCACCAATTAATGTTGAGATAGTAGTAATTAAGATGAAAGA 954
DB 3743 TGGCATCGATTACCTTTATGAAAAAACCCAGAAAATTTACGGTTTGAACGTATAA 3802
QY 955 TGA 957

```


/note="Similar to Escherichia coli hypothetical protein
yqcr SW:YQCR_ECOLI (P76010) (244 aa) fasta scores: E(): 0,
40.7% id in 241 aa"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="CAC92455.1"
/db_xref="GI:15981156"
/translation="MGPVSETSKQFVKKKLAICAILRLDQKNDTAVMVTHARGOFI
SLIDIQPTNFIIDFGSVENENVALGAEQITIVTEPTGAKIEFCNKLKHVEYLS
LPFSAIPDOLYFORREYFRVSIQWPAYVCSGKFPDGTQYKYTLADISLGMGLY
AKGSEFPLOGSVLRDAADVLAGFGLKLDLQFIRALDKQVNVNKGEMLTQRLSEFK
FLRSLIQEKGLORALFELEKQQTAKARFQSEL"
complement(4554. .5813)
/gene="proA"
/note="proA"
/complement(4554. .5813)
/gene="proA"
/EC_number="1.2.1.41"
/note="Similar to Serratia marcescens gamma-glutamyl
phosphate reductase ProA SW:PROA_SERMA (P17857) (417 aa)
fasta scores: E(): 0.85, 1% id in 417 aa, and to
Escherichia coli gamma-glutamyl phosphate reductase ProA
SW:PROA_ECOLI (P07004) (417 aa) fasta scores: E(): 0,
72.4% id in 417 aa"
/codon_start=1
/transl_table=11
/product="gamma-glutamyl phosphate reductase"
/protein_id="CAC92456.1"
/db_xref="GI:15981157"
/translation="MNLEHMGKAAQASWOLAMISTAKKNQALAVIANLLESOTI
LQNERDMAARESGEALDRLLTPARLAAIANDVQRCVNDPVRVIGESLSDI
SGKLERRVPLVIGVIEYARPNTVIDASCLKTGNAILRGKETHETNGATVNV
IQALBQQCLGPAAYVIESPDQVNEILLDRVDMLIPRGASLHKLCREQSTIP
VIRGEGVCHTEVDENAFKALVIENAKIQRPSACNSLETLLVHQAVAKRTFLPLLS
ARMHAFGVTLHASPLAMPYLADGKAKVAVAEADYDDEWLSDLNVDIVTDIDAADH
IREHGTSHSDAITRLSLSHAEYFVRVADSSAVVYNASTRTDGGQFGLGAEEAVSTQK
LHARGPWGLDALTYYKIGVGGDLVRS"
complement(4572. .5744)
/gene="proA"
/note="Pfam match to entry PF00171 aldedh, Aldehyde
dehydrogenase family, score -217.90, E-value 0.00013"
complement(4776. .4841)
/gene="proA"
/note="P501223 Gamma-glutamyl phosphate reductase
signature."
complement(5823. .6926)
/gene="proB"
/note="YPO3222"
complement(5823. .6926)
/gene="proB"
/EC_number="2.7.2.11"
/note="Similar to Serratia marcescens glutamate 5-kinase
ProB SW:PROB_SERMA (P17856) (367 aa) fasta scores: E(): 0,
92.9% id in 367 aa, and to Escherichia coli glutamate
5-kinase ProB SW:PROB_ECOLI (P07005) (367 aa) fasta
scores: E(): 0, 87.1% id in 365 aa"
/codon_start=1
/transl_table=11

Query Match 13.9%; Score 168.6; DB 1; Length 220050;
Best Local Similarity 58.4%; Pred. No. 5.7e-29;
Matches 294; Conservative 0; Mismatches 209; Indels 0; Gaps 0;
QY 434 ATATATTTATGGACAAATTCCTTGATGGACAATCTGATACAGAGAAATTCGTGGACAA 493
DB 171890 ATATATCGTCTGGACCCCAATCGATCACCCGACAGAGATCGCGAACACAGCAGCGTCA 171831
QY 494 AAAGGCTGTGCTGATGATAACAGGACTCAGTGGTTCAGGGAAGACTCTCTTCGATGCT 553
DB 171830 CAGGCGGTAGTCTTGTGGTTACCGGCTGTGGGTTCCGGGAATCCACTCTGGCCGT 171771
QY 554 GCACGTAGTCTGAGTTGCAATTCGAGAGGCCACCTCAGCTATGCTATGATGTTGAGACAC 613

DB 171770 GCCTCGAACAAGCGCTTTTCGCTCGCGGTGCAGTACTTACTTGGATGGTGATAAC 171711
QY 614 CTCAGACATGGCCTAAATAGAGATTTAAGCTTTTAAGCGACAAGACCGTGCAGAAAATATA 673
DB 171710 GTGCGCCATGGCCTCTGCCGCGATTTTGGCTTTTCTGACCGCTGATCGTGAATAATC 171651
QY 674 CGAAGATTTGGTGAAGTGGCAAAAGCTTTTTCGTGATGCTGGTGTCTATGATGATTCGTAGC 733
DB 171650 CGCGGGTTCGGGAGGTGGCTTAAACTGATGGTTGATGCCGGCTCTGGTGTTCACGGCG 171591
QY 734 TTGATATCTCCATACAGGAGAGATCGTGATGCGCATGCGCTGCTCTACTTTCCACATTTCTAAC 793
DB 171590 TTTATTTCCCTCATCGGCTGAACGGAATAATGATAAGATATGCTGGCATCGGTCAG 171531
QY 794 TTTATTTGAAGTATTTATTTGATTTGCCCTTAAAAATTTTGAAGCTCGTGATCCTTAAAGGC 853
DB 171530 TTTATTTGAGTATTTTGTGATACCGCTTAGCAATTTGTGAGCGCGTGAGCCGAAAGGC 171471
QY 854 CTATACAAGCTTCGACGTACAGGAAGATTAAAGGTTTTCACGTGGAATTCGATGATCCATAC 913
DB 171470 TTATACAAAAAGCCCGTCCCGAGAGCTTAAAAATTTTACAGGATCGATTCAGTTTAT 171411
QY 914 GAACCAACCAATTAATGGTGAGAT 936
DB 171410 GAATCGCTCGCTCTCCTGATAT 171388
RESULT 10
AB013390 51860 bp DNA Linear PLN 27-DEC-2000
LOCUS Arabidopsis thaliana genomic DNA, chromosome 5, TAC clone:K919.
DEFINITION AB013390 BA000015
ACCESSION AB013390.1 GI:3128137
VERSION AB013390.1
KEYWORDS
SOURCE Arabidopsis thaliana (strain:Columbia) DNA, clone_lib:Mitsui TAC
clone:K919.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (sites) Kotani,H., Nakamura,Y., Sato,S., Asamizu,E., Kaneko,T., Miyajima,N.
and Tabata,S.
Structural analysis of Arabidopsis thaliana chromosome 5. VI.
Sequence features of the regions of 1,367,185 bp covered by 19
physically assigned P1 and TAC clones
DNA Res. 5 (3), 203-216 (1998)
98403884
2 (bases 1 to 51860)
Nakamura,Y.
Direct Submission
Submitted (06-MAY-1998) Yasukazu Nakamura, Kazusa DNA Research
Institute, Department of Plant Gene Research; 1532-3, Yana,
Kisarazu, Chiba 292-0812, Japan (E-mail:ynakamu@kazusa.or.jp,
Tel:81-438-52-3935, Fax:81-438-52-3934)
Address for correspondence: kaos@kazusa.or.jp
For the latest information on annotation of this clone, please see
http://www.kazusa.or.jp/kaos/cgi-bin/agd_graph.cgi?c=K919
Genes with similarity to proteins in the databases are described in
'product' or 'note' qualifiers. Genes that have no significant
protein similarity are described as 'unknown protein'.
The software programs used to predict genes include: Grail
(Informatics Group, Oak Ridge National Laboratory,
http://compbio.ornl.gov/grail-1.3/),
GENSCAN (Chris Burge, MIT, http://CCR-081.mit.edu/GENSCAN.html),
NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of
Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and
SplicePredictor (Volker Brendel, Stanford University,
http://grenlini.zool.tastate.edu/cgi-bin/sp.cgi).
Genes encoding tRNAs are predicted by tRNAscan-SE
(Sean Eddy, Washington University School of Medicine, St. Louis,
http://genome.wustl.edu/eddy/tRNAscan-SE/).

This sequence may not be the entire insert of this clone. It may be shorter because we remove overlaps between neighboring submissions. The 5' clone is K8K14 and the 3' clone is LA522.

FEATURES

Location/Qualifiers	1..51860
organism="Arabidopsis thaliana"	
strain="Columbia"	
db_xref="taxon:3702"	
chromosome="5"	
clone="K919"	
clone_lib="Mitsui TAC"	
338..912	
/note="codon recognized: CAU; gene_id:K919.1"	
/product="tRNA-Met"	
/evidence=not_experimental	
complement(join(1150..1181,1265..2323,2389..2440))	
/note="gblAAD30228.1	
gene_id:K919.2	
similar to unknown protein"	
/codon_start=1	
/evidence=not_experimental	
/protein_id="BAB08454.1"	
/db_xref="GI:9757867"	
/translation="MRHLDYIFLFIAPLSFAGAGQESTPIEALNNLNNLLPOTDTDL DLVSVGNKITEISLIIKAEIWLTHILSRYPSTKITTIIVIFFPSCQTOHSSD LIVSLKNIYHLSLRWGLNIIKVSQFSYQCLNNPKSEMFKPVLIFLKTINSTFI NPQNFLTSPHNRDLHSVEKLAGSLGFKNVFNLPNPEQESTTMTNRNLSRLNLS SKFTSPPTLSPSPETSPHSSIGSPSPPTIPYFPSPQSPMESNOGISLPCLPYH PAPSPQVKKNVGELWCVAKPSVAATLQOISLDFACGGQGANCDIKPHGICYPDT YMASVAFNSYWKTKRNGKSCFSGGTAMLITDPSYQHCRFVLS"	
join(6001..7466,7558..7757,8265..8506,8587..9378)	
/note="gene_id:K919.3"	
/codon_start=1	
/evidence=not_experimental	
/product="formin-like protein"	
/protein_id="BAB08455.1"	
/db_xref="GI:9757868"	
/translation="MKALQSRFEFFFFFYIFFSVSSSEAHRRILHQLFPESPTPP PDFSSTPLPDTDPQFPFENPSTPQOTLFFPPPPVADVNGGLPIPTATQSAK PKKVAIVISGIVITLGLMSALAFPLVGRKAKHASDQTKLVTGGDGGGSRFOEDSG PPTTSTFIYMGTVETRVASASENGGTGPNVSSPYRKLNSAKRSRYRPSPELOP LPLHAKTPQSPNSPALSPPSSSGSECRDTAFYPHGSAISSDDGYVTAFFPSAG SIPHSKTSRSPKSGAPTAAASRPMKHVITPSIKQLPPVQPPPLGLESDGEG LPYQNKPKSQPPPPPNRAAFQALIOEKSPVPPRSPPLQTPPPPPPPPLAPP PPQKRPDFMLKRVNSEATNTTSPSRQAFKTPSPKTKAVEEVNSVVSAGLEKS GGDTDPSPKPLKPLHMDKVRASSDRATVMDLKSSSFQINEDMEHLFGCNSSSAP KEPVRSVIPLAENENRVLDPKKSQNTAILLRALNVTRIEVSEALTDGNPSLGAELL ETLVKMAPTKEEIKLREYSGDVSLGTAERFLKTLIDIPFAFKRVEAMLYRANFDAE VKYLRNSFOTLEASLEAKSRLFLKLEAVLMTGNRMNVNTRGDAFAFKLDTLLKL VDIKGVGKTTLLHFVYQBITRSEGTITTDEITLHGNNDGFRKQGVAGLSRDLY NVKKSAGMDVLSSVYTKLEMLDGLKSLFKTTTQGRFDSMTFLKEAEFEIRKI KGERKALSMVKEVTRYFHGNAEEAHPLEPVMVVRDFGLVDNVCKEVTQMENST SMGSASRSPRISATSLPVLHYKAKQDPTSDSEHSSNST"	
complement(join(10221..10664,10740..11046,11155..11285, 11426..11662))	
/note="gene_id:K919.4	
pir T04718	
strong similarity to unknown protein"	
/codon_start=1	
/evidence=not_experimental	
/protein_id="BAB08456.1"	
/db_xref="GI:9757869"	
/translation="WVTGCVDLHQFSKASDSSVPIPPPLPSKSDGLKKKLGHSSVST ATRMWRDLRNDGFKADVVYTDNGSIYIAHANILGTASTVTKGMLQAKRRHKWHTI SIRGVHADVAVRFIRFLSSCYKEEMNEFIMHLLLSHAYVVPOLKRVCSWHLHGL TLTVNVYDVLQALLCDFPLSLISHRMIMKHNELNLSATEAMTKKSHPLKEVERD SVITEAMTERMRKRDQRIYSQLYEAMEALVHICRDGCKTIGPHDKDFPNHATCN YEACKGLESILRHFACKLRVPGCVCKRMMWOLLELHRSVCAGSDQCRVPLCRNLKE KWEQSKKDSRWKLIVKNVLGSKKIGSPFFLPVINC"	
complement(join(13074..13266,13326..13506,13554..13860))	
/note="gene_id:K919.5	
unknown protein"	
/codon_start=1	
/evidence=not_experimental	
/protein_id="BAB08461.1"	
/db_xref="GI:9757870"	
/translation="MATNIVRLIAPSRVSSRFLPVRSRFLSSGTPPPPOAPSPNOD LNRDVQDQNLQOMKEEEGEGGGEFVNEEDTGEIGGPRGPPTRYGDMQR ACSYIFLGRMEINSFYVMYLNQNGEGRKCSRTSNLSPLFGSSFGDLVTPVTKNF FLRFSDLELRKQEMSLTNLVQSSGLTLFGGYMSSGSPSLGMAAWVSEMSRYFCCS TTLOCALL"	
join(14389..14437,14667..14734,15314..15424,15512..15598, 15690..15898,15983..16289)	
/note="gene_id:K919.6"	
/codon_start=1	
/evidence=not_experimental	
/product="porin-like protein"	
/protein_id="BAB08458.1"	
/db_xref="GI:9757871"	
/translation="MSKPGLEFDIGKAKDLITRDVNSDKFSISTYSASGVALTST ALKKGGVHAADVATQYKYKNALFDVKIDTSSVLTVTTEILPSTKAISFKVPDYN SAKLEVOYFHDHATVTAANAALKNPLIDITATLGSPIVIFSGAEAGDYDTKTFTKYNA GISVTKPDACLIIILGDGDSLKASYLHHFDEFKRTAAVGEVYRKFTSNEINTITVGL YALDHSTAVKAKLNNHGTILGALLQHEVLPRSLTVTSSEIDTKALEKHPRGLSLAKP "	
complement(17136..17576)	
/note="gene_id:K919.7"	
/codon_start=1	
/evidence=not_experimental	
/product="60S ribosomal protein L26"	
/protein_id="BAB08459.1"	
/db_xref="GI:9757872"	
/translation="MKPNPVRSSRRKRNKAHFTAPSSVRRVLSSPLSKDLRNKHNV RSMPTKDDVOVVRGTFKRGREKVMQVYRRKVIHIERITREKYNSTVNVGVNASN VMITKLRLDKRKLTERKANGRAADKEGTFESADYEMVND"	
join(18523..18736,18919..19050,19138..19406,19478..19562, 19698..19766,19905..20008)	
/note="gene_id:K919.8"	
/codon_start=1	
/evidence=not_experimental	
/product="adenylylsulfate kinase-like protein"	
/protein_id="BAB08460.1"	
/db_xref="GI:9757873"	
/translation="MARCVGRVYPAFGESESHLSERRFLKLSSTNSDPSAGSKSL KLRGCIHRMSYFRPIMAKDESISRSOGETKQINGKOKNLVHDCPTVKSORDELKO KCVITWTLGSSGSKSLACALSRALHNRGLSYLDGDNVHGLNSLSPENDRAE NIRRVGVALFDAGSLIGIASLISPYRIERACRALLPGGDFIEFVMDVPLHVCEAR DPKGLIKRAGAKIKGFTGVDDPYEAPLDCEVHIIISFSSSLCEMADIVVSYLDQN GYLKKHSTKGRNCC"	
join(20426..20553,20633..20695,20778..20881,20965..21056, 21129..21218,21301..21376,21472..21830,21907..22163, 22256..22544,22620..22730,22826..23044)	
/note="gene_id:K919.9"	
/codon_start=1	
/evidence=not_experimental	
/product="peptidylprolyl isomerase (cyclophilin)-like"	
/protein_id="BAB08461.1"	
/db_xref="GI:9757874"	
/translation="MGKQHSKDRMFIITKTWATEWGAKSKEKNTFPFKSLPYCCAL TFLPEDPVCTIDSGFEITITVIRKFGKHPVTGAPLGEDLIPLFHKNSEGEYH CPVLNKVFTEFTHIVAKTGNVFCYEAIKELNIKTNNKELLTEEPTRADITION PNAVDGKTVTFDHFVNGKLIDDEELKMNSDPAYNINSGDIKMLADLGTAKAEI ALHGGGNKARNERAAIAAILLESKIKEYSKAEQKQYTSVDAASAVSFGSADA AKAGSDKTAARTAMHAGDRTPVNSKMWKVSAGASRSFTSSAFTVTKNDPELLI KVERNPKKGVQFQTHGDLNLELHCDIAPRACENFILTLCERGVYNGVAFHSIRNF MIQGDPTKHTGKGESIGWKPFKDEPNKSLHLSGRGVSVNANGSGHTNGSQFFVLKYS ATHNLYRTVFGVVGGLATLAAMENVPVDESRLPEETIKLIEASVFNPTDEDEE EKEAKAEKNEKDIEKIGSWSNFGSGTTEAGAGGGGVGKYLKAMSSTATDKTGS LSDSTIGVSKKRTKTASATGFKDFSSW"	
complement(join(23419..24249,24339..24923))	
/note="emb CAB66926.1	
gene_id:K919.10	
strong similarity to unknown protein"	
/codon_start=1	
/evidence=not_experimental	
/protein_id="BAB08462.1"	

/d_b_xref="GI:9757875"
 /translation="MKNNKYNNKSTSLHCNDAGGCRSLTITVTVVGVGFVAHLIS
 LKPKNNHHQVSDQVLLVHHLAIPVRELIRVEEVLRMPKPKSPRTSKRSR
 KPIPLEEFLDKSDIRHLEFFGDKPTAAFGPTKDMGNETSYPPGKIMWDTQGNQGLV
 HGGGILDLVKSNTYKMDGEGDPTHAHKGPAPVDICGVCSKDLTKWKNQGLV
 LGAEYNTKDLHKSNSLRPKVINYNEKTEKTVYSMMVSHLDIADNTKASGVASINSNP
 PFYLYSKRPHGDSRMTVFKDDQVAGLYEYSMEVHDLPGLTEDYLDVTPVKR
 VMVQGHREAPALFKHQNITYMTVSTGWAPNEALHAARESIMPWEKLVNCPICGGNK
 VFLRTITFAQSYVITPLPGVPVGAEIFMADRNWPAADLRSSVWLVPLTGGSDQPLEF
 NFGFSSPSLSTVYHSHSKRWLP"
 complement(join(26074) . 26457, 26551 . 26727, 27046 . 27117,
 cds

RESULT 11	
BSY09476/c	
LOCUS	BSY09476 5353 bp DNA
DEFINITION	B-subtilisin 54kb genomic DNA fragment.
ACCESSION	Y09476
VERSION	Y09476.1 GI:2145361
KEYWORDS	addA gene; argC gene; argJ gene; deqA gene; lpi gene; nprB gene;
	linear BCT 31-MAY-1997

SOURCE ORGANISM	REFERENCE AUTHORS	TITLE
Bacillus subtilis.		
Bacillus subtilis.		
Bacteria, Firmicutes; Bacillus/Clostridium group;		
Bacillus/Staphylococcus group; Bacillus.		
1 (bases 43051 to 53533)		
Medina, N., Vannier, F., Roche, B., Autret, S., Levine, A. and		
Seror, S. J.		
Sequencing of regions downstream of addA (98 degrees) and citG (289		
degrees) in Bacillus subtilis		
Microbiology 143 (Pt 10), 3305-3308 (1997)		
98015415		
2 (bases 1 to 17287)		
Medina, N., Levine, A., Roche, B., Vannier, F. and Seror, S. J.		
A 17 kb nucleotide sequence downstream of addAB of the Bacillus		
subtilis chromosome		
Unpublished		
3 (bases 17288 to 43091)		
Roche, B., Autret, S., Levine, A., Vannier, F., Medina, N. and		
Seror, S. J.		

TITLE A Bacillus subtilis chromosome segment at the 100 degrees to 102 degrees position encoding 11 membrane proteins
JOURNAL Microbiology 143 (Pt 10), 3309-3312 (1997)
MEDLINE 98015416
REFERENCE 4 (bases 1 to 53533)
AUTHORS Seror,S.J.
TITLE Cloning and sequencing of 53 kbp from B. subtilis chromosome
JOURNAL Unpublished
REFERENCE 5 (bases 1 to 53533)
AUTHORS Seror S.J.
TITLE Direct Submission
JOURNAL Submitted (14-NOV-1996) Seror S.J., Universite Paris XI - CNRS, Institut de genetique et Microbiologie, Bat. 409, Unlversite Paris XI, Orsay 91405 cedex, FRANCE
COMMENT this sequence includes and completes 279580.
FEATURES Location/Qualifiers
source 1..53533
organism="Bacillus subtilis"
db_xref="taxon:1423"
<1..702
/codon_start=1
/transl_table=11
/product="Adda"
/protein_id="CAA70668.1"
/db_xref="GI:2145362"
/db_xref="SWISS-PROT:P23478"
/translotion="WTYPHQEVYQIRTKQSVSEIKRREYDEYSGRAPVKPDGSLYRRFAFMKKGFLTAAGKQVTHMHQIPLSHVPSIEAEQTVHRIKELLTEQKQIDAIDEETVQFPHTEIGQGLIGAKWKDRIDFSLALPAKEIYPDAAHEADPELLVQGIIDALYEDGLVLLDYKSDRIETGKFGHQFGAAGPILKKRYETQILYTRAVEQIARKTVKGCALYEDFGDGHILTL"
774..1514
note="difference with the nucleotide sequence previously described: one more G at position 1332"
/codon_start=1
/transl_table=11
/product="Sbcb"
/protein_id="CAA70669.1"
/db_xref="GI:2145363"
/db_xref="SPTREMBL:O06711"
/translotion="MRILHTADHLGKTLLEGSRSLSEQADVLDLELNTTVKDEQTDIAVMAGDAFTVNPALASQALFESLALSDRGKRPVTVIAGNHDPRLSAASPTLHNGIHLGYSPTTPIHEVPSAGELLAVGALVPSEARNEVLSDTFDEKLLRDHYDVKIRQAFEMHTSPRTDVAKIAASHLYVAGGNQTDSERPTEVGAVTVAESLPADAALRCA RPFASPANDQAGADACALPRISARLOLF"
1947..2816
note="putative"
/codon_start=1
/transl_table=11
/product="YirY"
/protein_id="CAA70670.1"
/db_xref="GI:2145364"
/db_xref="SPTREMBL:O06712"
/translotion="MKPIALSTKGLHGSFREQITDFEGLSGAGVFGIFGPTGSGCKSSI LDMATLYCKVRANNTNGHILNHAETLSVSFTALONTHQISYKRVKRYKRTDEM KVKTALCFTEIDETVHTLADKASPVANKRVEELGLTIDDTFRAVVLPGQKFAEFLSL KGARRHRLQRLNTEQYGDRLVYKKLRQQAENARKNEKLAEOGSLGEASAEAVQQA EKVLQEAEVRLAEWRKNRQQAERLTERHQELIWNVQEKSTYEEEEKRLAEAEQPHLDSM QKRLLEAETSSSP"
3010..4626
note="putative"
/codon_start=1
/transl_table=11
/product="YirZ"
/protein_id="CAA70671.1"
/db_xref="GI:2145365"
/db_xref="SPTREMBL:O06713"
/translotion="MQEIIKLSKQAEERKKADLRQKEALQSVMNELETVTDRLT RGNQRTQLKQDKLSQRTVSDERKSCQEEERKALIRQTEQIKRKRSESLNLYLO KMNKEKNTLVQTAEBENNIIOAYEAVQTVYHVLVCETSRKLTMRSEARKSOHTLHLQ REKARVALLTKELAQTLGAKGPCVCSOTDHPDSAVSHETIEDSHLEDLKRITDVLQ TEAALLSQEILSLKMLEQSAFTEQCPFIQITQANLEAASAEFHQPTVYAEFATLQ FEMWRIKQDILSVKTRMEQASIGAYQESLKKAEQNEKIGFKREKREDEHYSLELSQAS

```
CDS
MDSLLNMFKEAFQNSVDEAEKWOQAEIKDRAEECEKRIEKSIAFLABHEAQKEKL
RESGHLEKELELHYAERIKSVIADYEBELGDYAKGDSIPIQLRSVQODLKLKEK
EQLSYEEQSQMKNQAKSRASASELTIOEAKGRLEKAKAAWLEHTKNTSITRTEEV
EQSLIPADELEKMKGTIDQFMDKLNKMLQT"
4824..5339
/note="putative - Homology with SbcC from C. perfringens"
/codon_start=1
/transl_table=11
/product="Yisa"
/protein_id="CAA70672.1"
/db_xref="GI:2145366"
/translation="MHIDRLDKIQAFAVKGNTVFEEFLAEQLESVARDASARLSMLTRQ
RYAIEVDSGGFVMRDANGVRPVSSLSGGTEFTLTSIALALSQIOLRGEVPLQ
FFPLDGGFTLDLDDLTVTVALEKQLQSDNLAVGVISHVQELRLPKKLIIVHPAES
GRTRVSLRLM"
5353..5655
/note="putative"
/codon_start=1
/transl_table=11
/product="YisB"
/protein_id="CAA70673.1"
/db_xref="GI:2145367"
/translation="MAKQIAGICEIGRSDVQLTEHHLTPKEEGTFLPTAMLCIPCH
KQIHLYTNQELAVKLNGMAELRSDPELAFVWINKOPPEKLIKTRKSNERRKKK"
complement(5692..5919)
/note="putative"
/codon_start=1
/transl_table=11
/product="Yisc"
/protein_id="CAA70674.1"
/db_xref="GI:2145368"
/translation="MSFPAIVGPTAINISISGVNFGDSFYLSPKSSSKSALGSGAG
NTGDFLLNNAVNTNYIDPDVNDQDMVNG"
complement(5942..6343)
/note="putative"
/codon_start=1
/transl_table=11
/product="YisD"
/protein_id="CAA70675.1"
/db_xref="GI:2145369"
/translation="MLKRIKRILVKFNSLGIASVFQVQDNEIDMVKVFVAVORSLS
TFYHNSGKNKEYQIFQQAQVPLPETGVQSAFCHEVPAIYVRSIKIQGVASSVLH
AGSASLIRGDARLKHRIQSPRSQSPAKNI"
complement(6343..6519)
/note="putative"
/codon_start=1
/transl_table=11
/product="Yise"
/protein_id="CAA70676.1"
/db_xref="GI:2145370"
/translation="MIFTVINRSLEVDIRMNGVSSSVFHHGDTESILYLSSTFDTPP
ESLIIGPAPLAP"
complement(6516..7133)
/note="putative"
/codon_start=1
/transl_table=11
/product="YisF"
/protein_id="CAA70677.1"
/db_xref="GI:2145371"
/translation="WYDQSVSYLONLSNFVQQAIIHQQLERQLKEIOTEMNTMKQR
PARTIRVEYFDQLEKRLDGLNTGLNPTDFNSVQNFVDVSQSTPOLGMWQDEESAQ
LMQIQKNVMTLEETPEIDLEQNOYDRDDTNRRHVEDIRKQMDSRHHYMSH
IKKEENTPAYAEHIAEHVKRDIYRAVEHFLEHIPSEMKGDEQA"
complement(7156..7389)
/note="putative"
/codon_start=1
/transl_table=11
```

```
CDS
/product="YisG"
/protein_id="CAA70678.1"
/db_xref="GI:2145372"
/translation="MNFYINOTIOLNVLRESINSSILOGSAGSIKSLSNLYNTGS
YVEPAPEVSGSQPLQLOEPDGTGSLVPLQPPGR"
complement(7404..7625)
/note="putative"
/codon_start=1
/transl_table=11
/product="YisH"
/protein_id="CAA70679.1"
/db_xref="GI:2145373"
/translation="MPAIVGAEPKINAICTGVHIGDCIITISPOAQVRTFAGAGSFNT
GDSLKVMNYQNATNYVDNDAYDQPIVANA"
complement(8048..8216)
/note="putative"
/codon_start=1
/transl_table=11

Query Match      13.3%; Score 161.4; DB 1; Length 53533;
Best Local Similarity 57.1%; Pred. No. 2.2e-27;
Matches 294; Conservative 0; Mismatches 221; Indels 0; Gaps 0;

QY 432 CTAATATTTTATGGCACAATTCCTTGATTGGACAATCTCATAGACAGAAATTCGTGGGAC 491
Db 27663 CGAACATCATTTTGGCATCCCGCTGCCATCTCAAAGTCTGACAGACAGTCTCTTAACGGC 27604

QY 492 AAAAAGGCTGTGTCGTATGATAACAGGACTCAGTGGTTTCAGGAAAGTACTCTTCAT 551
Db 27603 ACAAAGCTGCGCTCTTGGTTTACAGTTTGTCCGGCTCCGGAAATTCGGTCTGGCCA 27544

QY 552 GTGCACTGAGTCGTGAGTTGCATTGCAGAGGCCACCTCAGCTATGTACTTGAATGGTGACA 611
Db 27543 ATGCTGTCATGAAAAGCTTTACCGCAAGGGCATTCAGAGCTACGTGCTTGACGGTGATA 27484

QY 612 ACCTCAGACATGGCCTAAATAGAGATTTAAAGCTTTAAAGCAGAAAGCGTCAGAAAATA 671
Db 27483 ATATCCGCCACGSCTTAAACAAGATCTTGATTTTCAGACTGTCACAGGATTCAAAACA 27424

QY 672 TAGCAGAGTTGGTGAAGTGGCAAGCTTTTTCGTGATGCTGCTGATGCTATGATGCTGCTA 731
Db 27423 TCCGCGGATCGGAGAGTGGCAAGCTGTTGTGACAGAGCGGCAAAATGATTTTAACGG 27364

QY 732 GCTTGATATCTCATACAGAGAGATCGTGATGCATGCGCTGCTCTACTTCCACATTTCTA 791
Db 27363 CTTTTATTCTCCATTCAGAGAGGATCGGCATGTTAGAGCGCTCTTCTCTAAAGTG 27304

QY 792 ACTTTATTGAAGTATTATTGATTGGCCCTAAAAATTTGTGAAGCTCGTGATCCTTAAG 851
Db 27303 AGTTCTTCGAAATTTATGTGAAGTGGCCGCTGCATGTTGTGAACAGCGGATCCAAAG 27244

QY 852 GCCTATACAAAGCTTCACGCTACAGAAAGATTAAGGTTTTCACGTGAATGATGATCCAT 911
Db 27243 GGCTTTATAAAAGCGGAAACGAGAGATCAACATTTACCGGGGATCGATTCCTCGT 27184

QY 912 ACGAACCAACAATTAATGCTGAGATAGTAATTAAG 946
Db 27183 ATGAAGCGCGCTTTTACCCGATTTTCATCATTCAG 27149

RESULT 12
BSUB0006/C      210440 bp      DNA      linear      BCT 26-NOV-1997
LOCUS      Bacillus subtilis complete genome (section 6 of 21): from 999501 to
DEFINITION      1209940.
ACCESSION      299109 AL009126
VERSION        299109.1 GI:2633260
KEYWORDS
SOURCE
ORGANISM      Bacillus subtilis.
               Bacillus subtilis.
               Bacteria; Firmicutes; Bacillus/Clostridium group;
```

Bacillus/Staphylococcus group: Bacillus.

1 (bases 1 to 210440)

REFERENCE AUTHORS

Kunst, F., Ogasawara, N., Moszer, I., Albertini, A.M., Alloni, G., Azevedo, V., Bertero, M.G., Bessieres, P., Bolotin, A., Borchert, S., Boriss, R., Boursier, L., Brans, A., Braun, M., Brignell, S.C., Bron, S., Brouillet, S., Bruschi, C.V., Caldwell, B., Capuano, V., Carter, N.M., Choi, S.K., Codani, J.J., Connerton, I.F., Cummings, N.J., Daniel, R.A., Denizot, F., Devine, K.M., Dusterhoft, A., Ehrlich, S.D., Emerson, P.T., Entian, K.D., Errington, J., Fabret, C., Ferrari, E., Galloway, D., Fritz, C., Fujita, M., Fujita, Y., Fuma, S., Galizzi, A., Galleron, N., Ghim, S.Y., Glaser, P., Goffeau, A., Gollightly, E.J., Henaut, G., Guiseppi, G., Guy, B.J., Haga, K., Haiech, J., Harwood, C.R., Henaut, A., Hiber, H., Holsappel, S., Hosono, S., Hullo, M.F., Itaya, M., Jones, L., Joris, B., Karamata, D., Kasahara, Y., Klaerr-Blanchard, M., Klein, C., Kobayashi, Y., Koetter, P., Koningsstein, G., Krogh, S., Kumano, M., Kurita, K., Lapidus, A., Lardinots, S., Lauber, J., Lazarevic, V., Lee, S.M., Levine, A., Liu, H., Masuda, S., Maue, C., Medique, C., Medina, N., Mellado, R.P., Mizuno, M., Mostl, D., Nakai, S., Noback, M., Noone, D., O'Reilly, M., Ogawa, K., Ogihara, A., Oudega, B., Park, S.H., Parro, V., Pohl, T.M., Portetelle, D., Porwollik, S., Prescott, A.M., Presecan, E., Pujic, P., Punelle, B., Rapoport, G., Rey, M., Reynolds, S., Rieger, M., Rivolta, C., Rocha, E., Roche, B., Rose, M., Sadaie, Y., Sato, T., Scanlan, E., Schleich, S., Schroeter, R., Scoffone, F., Sekiguchi, J., Sekowska, A., Seror, S.J., Serror, P., Shin, B.S., Soldo, B., Sorokin, A., Taccioni, E., Takagi, T., Takahashi, H., Takemaru, K., Takeuchi, M., Tamakoshi, A., Tanaka, T., Terpstra, P., Tognoni, A., Tosato, V., Uchiyama, S., Vandenbol, M., Vannier, F., Vassarotti, A., Viari, A., Wambutt, R., Wedler, E., Wedler, H., Weitzenecker, T., Winters, P., Wipat, A., Yamamoto, H., Yamane, K., Yasumoto, K., Yata, K., Yoshida, K., Yoshikawa, H.F., Zumbstein, E., Yoshikawa, H. and Danchin, A.

The complete genome sequence of the gram-positive bacterium

Bacillus subtilis

Nature 390 (6657), 249-256 (1997)

98044033

2 (bases 1 to 210440)

Kunst, F., Ogasawara, N., Yoshikawa, H. and Danchin, A.

REFERENCE AUTHORS

Submitted (18-NOV-1997) I. Moszer, A. Danchin, Institut Pasteur, Regulation de l'Expression Genetique, 28 rue du Docteur Roux, 75724 Paris Cedex 15, FRANCE. E-mail: moszer@pasteur.fr, adanchin@pasteur.fr Phone: +33 (0)1 45 68 84 41, Fax: +33 (0)1 45 68 89 48

FEATURES

source

1. .210440

/organism="Bacillus subtilis"

/strain="168"

/db_xref="taxon:1423"

369. .1721

/gene="yhxA"

369. .1721

/gene="yhxA"

/function="unknown"

/note="similar to

adenosylmethionine-8-amino-7-oxononanoate

aminotransferase"

/codon_start=1

/transl_table=11

/protein_id="CAB12765.1"

/db_xref="GI:2633261"

/translation="MEMMGNICQOGLKQKDEQVWAMKGAHQADSLIAQKAEGA

WYTDGRRYLDAMGLWCYNIGYGRKEIAEAYEQLKELPYPLTOSHAPAIQAEK
LNEWGGDYVIFFSNGSEANEAFKIAQYHLQNGHSRYKFLSYRAYHGNLTGAL
SATGQAKRYEYFSLSGFLHAAFPDPIYRNPDADLESADRIMFWELSETIAGV
IMEPIITGGILMPDGMKVEDICRRHALLICDEIVCGFTGRTGPFMGHYGVKP
DIITWAKGIT SAYLPLSATVKRDIPEAYGEAPDRFRHVNTEGSPACALAKNL
QMEDEQIQRSLRGLAKLIGLQALREHPAYGVDRGKGLLIGELVYLKTEKAPADA
AKVQNWAAKCEKGLIIGKNGDTRVAGYNNVHVAPPFCLEEDLSFIVTKVESFOTI
"

1749. .2441

/gene="glpP"

1749. .2441

gene

CDS

/gene="glpP"

/function="control of mRNA stability of glpD"

/codon_start=1

/transl_table=11

/product="transcription antiterminator"

/protein_id="CAB12766.1"

/db_xref="GI:2633262"

/db_xref="SWISS-PROT:P30300"

/translation="MMSFNQPIILPAIRNKQDFENLSSFSYGVILDIHLQKLGVI
KEAQHCKNMVHVDLQIKHDEYCAEFICQDIKPAGIISTRSNVIAKAKOKKIYAI
REOFLHTSAMKSMFEICKHKPDEFIEVLPGVPSLIQELKERTGPIPAFGGFIREE
DVEQALKAGAVATTNTNKLKKYENFLTESD"

2406. .2441

/gene="glpP"

2506. .3330

/gene="glpP"

2506. .3330

/gene="glpP"

/function="glycerol utilization"

/codon_start=1

/transl_table=11

/product="glycerol uptake facilitator"

/protein_id="CAB12767.1"

/db_xref="GI:2633263"

/db_xref="SWISS-PROT:P18156"

/translation="MTAFWGEVIGTMLLIIFGAGVCAGVNLKSLFSQSGWIVVFGW
GLGVMAAYAVGGISGAHLNPALTIALAFVGDGFWKQVPPVIAAQMGAIGAVIYL
HYLPHKSTDPAKGLGVSTGSPITPHFANVLSEVIGTGVVLGILAGANQFTGL
NPLIVGLIIVAGISIGTGTVAINPARDLGPRIAHAFLPDPCGSSNNKWKYAMVPVG
PILGSGFGVYNAAFKGHITSSFWIVSVILVVLGLLYVYTKSHSAKTLNSKYL"

3349. .4942

/gene="glpK"

3349. .4839

/gene="glpK"

/EC_number="2.7.1.30"

/function="glycerol utilization"

/codon_start=1

/transl_table=11

/product="glycerol kinase"

/protein_id="CAB12768.1"

/db_xref="GI:2633264"

/db_xref="SWISS-PROT:P18157"

/translation="METYILSDQGTSSRAILFNKKGKIVHSAOKEFTQYPPHGW
EHNANEIWSGLAVIASVISESGISAGIAGITNQRTTVMVDKDGSPYNAIVW
QSROTSGICELEKGYNDKFKETGLLDIPYFGTGVKWLNDNVEGAREKAEGLL
FGTIDTLWKMSSGKAHYTDYSNASRTLMENIYDLKMDLDDIIGVPKSMLPVKP
SSHVYAEVDYHFGKIDPIAGAGDQGSALFQOACEEGMGKNTYGTGCPMLMNTGE
KAIAKSEGLTTIANGIDKVNVALEGSIFVAGSAIQWLDGLRMFDQSSLSSEYAEK
VDSTDGVYVPAFVGLGTPYDSDVRSVGLTRGTTRKEHFIRATLESATYQTKDVL
AMEADSNISLKTURVDGAVKNNLMQFQDGLNVPVPERPEINETTALGAAYLAGI
GFWRDSEIANQWNLDRPEPELEEKREKELYKGWQKAVKAAKAFK"

4895. .4942

/gene="glpK"

4980. .6675

/gene="glpD"

4980. .6647

/gene="glpD"

/EC_number="1.1.99.5"

/function="glycerol utilization"

/codon_start=1

/transl_table=11

/product="glycerol-3-phosphate dehydrogenase"

/protein_id="CAB12769.1"

/db_xref="GI:2633265"

/db_xref="SWISS-PROT:P18158"

/translation="MNMHOFSLERDMLTDMTKKTYVDLFIIGGTTGATGATLDAASR
GMKYALSEMODFAAGTSSRSKTLKHGRLYLKQPEVMAEYKGERAIYVENGPHVT
PEWMLLPFGHGTGFTTISIGLYVDFLAGVKSERRSMLSAKETLQKPLVKDGL
KGGYVVEYRTDARLTIEVKEAVEFGAEPVYSKVELLYEKGAAGVLLVEDYTK
KEYKYAAKTIVNATGPWDLREKDKHNGKHQHTKGHLVFDQSVPLPKQAVYFDT
PDGMVFAIPRECKTVGTTDFVYKEALEHPMTEDRDYVJKSYNPEPLNITAND
IESWAGLRPLIHEEGKQDSELSRKDELSDSLITAGGKLTGYRKMADDIVDLVR
DLRAKEGKDFGPKCKTNMPSISGGHVGSNNKLSFVTAKTKBGTGAAGUSEKADQKLL

```
RYGSDVRVDFRVEALKDEAAKRNIPVHILAEAYISIEEMTATPAOFFVRRTRGLRF
DINWYTYDAYIDVDFMSERFOWDEQAKNKHTEINLKLHDAVVPLEQ"
6642..6675
/genes="g1pD"
terminator
6655..6672
/genes="g1pD"
terminator
6779..8512
/genes="yhxB"
gene
6779..8476
CDS
/function="unknown"
/notes="similar to phosphomannomutase"
/codon_start=1
/transl_table=11
/protein_id="CAB12770.1"
/db_xref="GI:2633266"
/db_xref="SWISS-PROT:P18159"
/transl_table=11
/translation="MTWRKSTERNKQTEHLDLELKERLIELEGDEQALEDCFYKDLFE
GTGMGEIGAGTNRMNITVVRKASGFAAYISKQEEAKKRGVVIAYDRKHSPEFA
MEAAKTLATQGIOTYVDFDELPPELSFAVRQLNAYGGVVVYASHNPPEYNGYKVG
DGGQLPPKADIVIBQVNAIENELTITVRYDKVTEKLISVHPPELSEVYKLVF
TPLHGTAANKPRRGLEALGYKNVTVVKQELPDSNFSTVSPPEEHAAYEAIKLG
EONADILIAATPDADRGLIAVNDQGRYVLTGNOTGALLLHYLLSEKKOGILPDNG
VLKTTVTSEIGRAVASSFGDLTDITLGPKEIGEKIKEYRASGOYFOQYRESKY
LIGDFARDKDALQALLAVENCAFYKQGSLSYEALINLFNEYGFYREGLKSLITLKG
QGAEQILALISAFQNPQKRMAGQGVVTAEDYAVSRKTLITSEKEADILPKSNILKY
FLEGGSWFCLRPSPGTEPKFYFAVFGSSLEDSERKLAVLSDEDVKNKTVDEIVESTAK"
8488..8512
terminator
/genes="yhxB"
gene
8625..9764
/genes="yhcy"
CDS
8625..9764
/genes="yhcy"
/function="unknown"
/notes="similar to two-component sensor histidine kinase
[yhcZ]"
/codon_start=1
/transl_table=11
/protein_id="CAB12771.1"
/db_xref="GI:2633267"
/db_xref="SPTREMBL:O07527"
/transl_table=11
/translation="MSKTRMEKLTAKTIAETLNQGHDIKATLDEVLKELSLTNLQS
GWFLIEEGSTLAADAYILPALSRRKRVLMCEGCYCLTKFNNGGILRAANIMCK
RISAEMLHCFDEGTGTHATVPLEDDRRFLGVAAAGKTMFDEDEHLHLESVAFQ
IGTAIORMLSLEYQOKNALLMERNRLAEQLHDSVNQMLFSVSLTAKAKTLTKDENIQ
Query Match 13.3%; Score 161.4; DB 1; Length 210440;
Best Local Similarity 57.1%; Pred. No. 2.9e-27;
Matches 294; Conservative 0; Mismatches 221; Indels 0; Gaps 0;
Qy 432 CTAAATATTTATGGCACAANTGCTTGAATGGACATCTGATAGACAGAAATGCTGGGAC 491
Db 170437 CGAACATCATTTGGCATCCCGCTGCCATCTCAAAGTCTGCAGACAGATCTTAACGGAC 170378
Qy 492 AAAAAAGCTGTGCTATGATTAACAGGACTCAGTGGTTCACGGAAAGTACTCTTGGCAT 551
Db 170377 ACAAAGCTGCTCTTGTGTTACAGGTTTGTCCGGCTCGGAAATCGGCTGGCCA 170318
Qy 552 GTGCACATGAGTGTGATTCGATTCAGAGAGCCACCTCAGTATGATCTGATGTTGACA 611
Db 170317 ATGCTGTGATGAAGAGCTTTACCGCAAGGCGATTCAGAGCTAGCTGCTTGACGGTGATA 170258
Qy 612 ACCTCAGACATGGCTAAATAGATTTTAAGCTTTAAGCGCAGAGACCGCTGCAGAAAATA 671
Db 170257 ATATCGCCACGGCTTAACAAAGATCTTGGATTTACAGCTGGTGACAGGATGAAACA 170198
Qy 672 TAGCAAGAGTGGTGAAGTGGCAAGCTTTTGTGCTGATGCTGGTGTGCATATGCATGCTA 731
Db 170197 TCCGCGGATCGGAGAGTGGCAAGCTTTTGTGCGACAGCGGCAAGATGATTTAAACGG 170138
Qy 732 GCTTGATATCTCCATACAGGAGATCGTGATCGATCGCGTCTACTTCCACATTTCTA 791
Db 170137 CTTTATTTCTCCATTCAGAGAGGATCGGACATGTTAGAGCGCTCTTCTCTAAAGTG 170078
```

```
Qy 792 ACTTATTGAAGTATTATTGATTTGCCCTTAAAAATTTGTGAAGCTGTCATCTAAAG 851
Db 170077 AGTTCTTCGAAATTTATGTGAAGTCCCGCTGCATGTTTGTGAACAGCGCATCAAAG 170018
Qy 852 GCCTATACAAGCTTCAGCTACAGAAAAGATTAAAGTTTCTACTGGAATTCATGATCCAT 911
Db 170017 GGCTTTATAAAAAGCGGAAACGAGAGATCAACATTTACCGGATCGATCTCCGT 169958
Qy 912 ACGAACCAACAATTAATGGTGAGATAGTAATAAG 946
Db 169957 ATGAAGCGCGCTTTCACCCGATTTTCATCATTTGAG 169923
```

```
RESULT 13
AC104429
LOCUS AC104429 154330 bp DNA linear. HTG 11-DEC-2001
DEFINITION Oryza sativa chromosome 3 clone OSJNB00006P09, *** SEQUENCING IN
PROGRESS ***, 8 ordered pieces.
AC104429
AC104429.1 GI:17488564
VERSION HTG: HTGS_PHASE2.
KEYWORDS Oryza sativa.
SOURCE Oryza sativa.
```

```
ORGANISM
Oryza sativa.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
```

```
REFERENCE
AUTHORS Wang,R.A., Yu,Y., Soderlund,C., Chen,M., Kim,H.-R., Rambo,T.,
Saski,C., Henry,D., Oates,R. and Simmons,J.
1 (bases 1 to 154330)
```

```
TITLE
JOURNAL Rice Genomic Sequence
REFERENCE
AUTHORS Unpublished
```

```
TITLE
JOURNAL 2 (bases 1 to 154330)
Wang,R.A., Yu,Y., Soderlund,C., Chen,M., Kim,H.-R., Rambo,T.,
Saski,C., Henry,D., Oates,R. and Simmons,J.
Direct Submission
```

```
COMMENT
Submitted (11-DEC-2001) Clemson University Genomics Institute,
Clemson University, 100 Jordan Hall, Clemson, SC 29634, USA
* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* been provided by the submitter.
```

```
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
```

```
* 1 1981: contig of 1981 bp in length
* 1982 32672: contig of 30691 bp in length
* 32673 87321: contig of 54649 bp in length
* 87322 94650: contig of 7329 bp in length
* 94651 96734: contig of 2084 bp in length
* 96735 97792: contig of 1058 bp in length
* 97793 139678: contig of 41886 bp in length
* 139679 154330: contig of 14652 bp in length.
```

```
FEATURES
Location/Qualifiers
1..154330
/organism="Oryza sativa"
/db_xref="taxon:4530"
/chromosome="3"
/cclone="OSJNB00006P09"
```

```
BASE COUNT 43409 a 33401 c 33796 g 43714 t 10 others
ORIGIN
```

```
Query Match 12.9%; Score 156.8; DB 2; Length 154330;
Best Local Similarity 71.0%; Pred. No. 3.4e-26;
```


/gene="ci-Asak"
/codon_start=1
/product="ATP sulfurylase/APS kinase"
/protein_id="BAB00629.1"
/db_xref="GI:9229920"
/translation="MASATANHGGDDIOEKPPMSSNIVYQOHHSVSRDERGQVMCKGAER
GCTVFTGLSGAGSTLSMKLEIEILCSKGIPIAYSLDGNIRHGLNKDLGFAPADREEN
IRRGVAKLFADAGVVCVLSFISPYRKQSRARHNKSNLPEFIEVDFDPLQVCEG
RDVGLYKARQGGIKGTGIDSPYEPENAEVVTADMEVEDCMLVVKELMERNI
IPCEITLPVMEIFTSHSQVGKRAEALPSDITKLDQWLQVISEGWATPLTGFMR
ERFELQCQHFGLLKNPINOQSVPIVPHVSADKERVENAEIALKYDQGVKAILHKP
EFYPLHKERSROWGTSNKGHPHIMESGDWLCGGDIEVLERITWGDGLDKYRMT
PLELRAKFSNMADAVAFOLRNPVNHGALLMODTKBKLVGERKPNVLLHPLGGW
TKSDDPVLDVRMKOHSVLEGVLDPOSTVVAIFPAPMYAGPTEVOWHAKBELATGA
NFYIVGRDPAGMPHDEKRLIYDHSGRVLVMAEGLGRFEIVFRVAAYNKTTKAMD
FYDPENHNDFFEISGTRMRRAREGESLPEGFMAPKAMDVLVYQSLPRAT"

BASE COUNT	714 a	493 c	549 g	622 t
ORIGIN				
Query Match	12.8%	Score 155.6;	DB 3;	Length 2378;
Best Local Similarity	60.1%	Pred. No. 2.6e-26;		
Matches 279;	Conservative 0;	Mismatches 179;	Indels 6;	Gaps 1;
Qy 497	GGCTGTGCTGATGGATAACAGACCTCAGTGGTTTCAGGGAAAAGTACTCTTGCATGTGCA	556		
Db 142	GGCTGCACCGTTTGGTTTACTGGGTTGTGGGCGCTGCTAAAGTACGTTGTCATGAAA	201		
Qy 557	CTGAGTCGTGATGTCAGAGGCCACCTCACGTATGTACTTGTGATGGTGACAAACCTC	616		
Db 202	TTGGAAGAGTATCTTTGCTCCAAGGATATCCAGCTTACAGTTTGGATGGTGCACACATA	261		
Qy 617	AGACATGCGCTAAATAGAGATTTAAGCTTAAAGCAGAGACCGTCAGAAAATATACGA	676		
Db 262	AGACATGCGCTGAACAAGGACCTCGGTTTCGCACAGCAGATCGTGAAGAGATATCCGA	321		
Qy 677	AGAGTTGGTGAAGTGGCAAGCTTTTTCGTGATGCTGGTGTCATATGCATGCTAGCTTG	736		
Db 322	AGAATTGGAGAAGTTGCAAGCTATTTCGCCGATGCTGGTGTAGTGTCTGCTCTCATTT	381		
Qy 737	ATATCTCCATACAGGAGATCGTGATCGCGTCTCTACTTCCACATTTAC--	793		
Db 382	ATCTCTCTTACAGGAGGACACAAAGTGCACGAGGGTGCACAAATCCAACTTA	441		
Qy 794	---TTTATGAAGTATTTATGATTTGGCCCTTAAAAATTTGTGAAGCTCGTATCCTAAA	850		
Db 442	CCATTCTTGAAGTCTTTGTGGACACACCACTGCAAGTGTGTGAGGGCGGACGTGAAA	501		
Qy 851	GGCTTATCAAGCTTGCACCTACAGGAAAGATTAAAGTTTCACTGGAAATTGATGATCA	910		
Db 502	GGCTTCTATAGNAAGCAGGAGGCCAAATTAAGGGTTCACGTGATTCCTCTCTCT	561		
Qy 911	TACGAACCAACCAATTAATGTGATAGTAGTAATTAAGATGAAGA	954		
Db 562	TACGAACCAACCAAAATGCTGAGTTGTACAAAAACGGCCGA	605		

Search completed: November 2, 2002, 00:28:42
Job time : 2702 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2002, 22:55:32 ; Search time 234 Seconds

(without alignments)
8929.423 Million cell updates/sec

Title: US-09-720-384A-3

Perfect score: 1217

Sequence: 1 gcgtcgtttcattcatca.....aaaaaaaaaaaaaaaaaaaaa 1217

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_032802:*

- 1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
- 2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
- 3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
- 4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
- 5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
- 6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
- 7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
- 8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
- 9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
- 10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
- 11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
- 12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
- 13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
- 14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
- 15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
- 16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
- 17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
- 18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
- 19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
- 20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
- 21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
- 22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
- 23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
- 24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1217	100.0	1217	21	Corn Adenyllysulph
2	359.6	29.5	928	21	Wheat Adenyllysulph
3	338.8	27.8	936	21	Soybean Adenyllysulph
4	320.4	26.3	627	21	Arabidopsis thalia
5	316.2	26.0	890	21	Corn Adenyllysulph
6	296.4	24.4	1068	21	Arabidopsis thalia
7	296.4	24.4	1175	21	Arabidopsis thalia
8	296.4	24.4	1183	21	Arabidopsis thalia
9	290	23.8	1130	21	Arabidopsis thalia

10	288.4	23.7	1133	21	AAC36627	Arabidopsis thalia
11	263.6	21.7	917	21	AAZ298337	A. thaliana gene i
12	263.6	21.7	948	21	AAC42294	Arabidopsis thalia
13	225	18.5	714	21	AAC51190	Arabidopsis thalia
14	198.2	16.3	483	21	AAC51315	Rice Adenyllysulph
15	197.2	16.2	431	21	AAC50161	Wheat Adenyllysulph
16	162.8	13.4	521	21	AAZ50164	Saccharomyces carl
17	156.4	12.9	1160	21	AAA08597	Human transferase
18	150.6	12.4	2617	21	AAZ94206	Human Adenosine 5'
19	145.8	12.0	582	21	AAC49624	Human Adenosine 5'
20	139	11.4	990	16	AAZ04849	Human APS kinase/A
21	139	11.4	1875	20	AAZ84897	Human disease rela
22	139	11.4	2507	19	AAV33482	Human secreted pro
23	139	11.4	2527	20	AAZ00717	Human secreted pro
24	139	11.4	2527	20	AAZ00686	Human DNA sequence
25	139	11.4	2646	24	AAZ94855	Human DNA sequence
26	130.2	10.7	2990	22	AAS33174	DNA encoding human
27	130	10.7	606	22	AAH81465	Escherichia coli p
28	106.2	8.7	347	22	AAH81282	Escherichia coli n
29	95	7.8	868	21	AAF07688	Fusarium venenatum
30	85.6	7.0	3332	23	AAZ88909	DNA encoding novel
31	82.6	6.8	447	21	AAC35996	Zea mays DNA fragm
32	81.2	6.7	195	22	ABA71730	Human foetal liver
33	81.2	6.7	195	22	ABA37826	Probe #16292 for g
34	81.2	6.7	195	22	AAK20089	Human brain expres
35	81.2	6.7	195	22	AAK46143	Human bone marrow
36	81.2	6.7	195	22	AAI52052	Probe #20738 used
37	74.4	6.1	650	21	AAF14521	Aspergillus oryzae
38	69.2	5.7	588	21	AAF11283	Aspergillus niger
39	69.2	5.7	600	16	AAZ04850	Human Adenosine 5'
40	68.6	5.6	471	22	ABA59209	Human foetal liver
41	68.6	5.6	471	22	ABA27965	Probe #6431 for ge
42	68.6	5.6	471	22	AAK07411	Human brain expres
43	68.6	5.6	471	22	AAK33189	Human bone marrow
44	68.6	5.6	471	22	AAI38986	Probe #7672 used t
45	60	4.9	349980	22	AAH41225	Pyrococcus abyssi

ALIGNMENTS

RESULT 1
AAZ50160
ID AAZ50160 standard; cDNA; 1217 BP.

AC AAZ50160;

DT 04-MAY-2000 (first entry)

DE: Corn Adenyllysulphate kinase-2 cDNA clone.

XX Adenyllysulphate kinase; Adenosine-5'-phosphosulphate kinase; APS kinase;
3'-Phospho-adenosine-5'-phosphosulphate; PAPS; sulphate assimilation;
KW corn; clone p0016.ctscj40rb; transgenic plant; screen; antibody; ss.

OS Zea mays.

XX Key Location/Qualifiers

FT CDS 2..1033

FT /tag= a

FT /product= "Corn APS kinase-2"

FT /note= "Derived from clone p0016.ctscj40rb"

XX W0200004165-A1.

XX 27-JAN-2000.

XX 13-JUL-1999; 99WO-US15809.

XX 14-JUL-1998; 98US-0092833.

XX (DUPO) DU PONT DE NEMOURS & CO E I.

PI Palco SC, Allen SM, Anderson SL;
XX WPI: 2000-182430/16.
DR P-PSDB; AAY44789.
XX
XX New nucleic acid molecule and chimeric gene encoding an adenosine-5'
PT phosphosulfate kinase, useful for altering expression of sulfate
PT assimilation protein in plants -
XX
XX Claim 3; Page 30; 42pp; English.
XX
XX The present sequence is a cDNA encoding corn adenylylsulphate kinase
CC (APS kinase), also known as adenosine-5' phosphosulphate kinase. This is
CC obtained from clone p0016.ctscj40rb, derived from corn pooled tassel
CC shoots, p0016 cDNA library. APS kinase is a sulphate assimilation
CC protein, that catalyses the conversion of adenosine-5' phosphosulphate
CC (APS) to 3'-phospho-adenosine-5' phosphosulphate (PAPS). This sequence
CC is used as probe and primers to identify, obtain and synthesise sulphate
CC assimilation proteins from other plants. It is also used to produce
CC transgenic plants, that are useful for altering the expression levels of
CC a sulphate assimilation protein. The APS kinase peptides are useful for
CC producing antibodies, that are used to screen and isolate cDNA clones.
XX
XX Sequence 1217 BP; 344 A; 292 C; 303 G; 278 T; 0 other;
SQ

Query Match 100.0%; Score 1217; DB 21; Length 1217;
Best Local Similarity 100.0%; Pred. No. 1.2e-296;
Matches 1217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGTCGGTTTCATTCATCAATCAACAGAACCTCTGTGCACACAGCAGCAACACC 60
Db 1 GGGTCGGTTTCATTCATCAATCAACAGAACCTCTGTGCACACAGCAGCAACACC 60
Qy 61 GAGCCAGCGCCCGCCAGCCAGCCAGCGCCAGCGCCAGCGCCAGCGCCAGCGCC 120
Db 61 GAGCCAGCGCCCGCCAGCCAGCCAGCGCCAGCGCCAGCGCCAGCGCCAGCGCC 120
Qy 121 GAGCGGAGCGTCCGCGTCACTCTGTAATCCACAGCGCGCGCTCCGCTCCAGG 180
Db 121 GAGCGGAGCGTCCGCGTCACTCTGTAATCCACAGCGCGCGCTCCGCTCCAGG 180
Qy 181 CCTCACCCCTAGCGATCGCCACTCCCGCGCTCGTGATCCATGGCTCACTCCCGTTC 240
Db 181 CCTCACCCCTAGCGATCGCCACTCCCGCGCTCGTGATCCATGGCTCACTCCCGTTC 240
Qy 241 CTCACACTCTTCCGGGGTCTCGCAGTATAGTGGCGCGCGAGGGGAGGCGCGG 300
Db 241 CTCACACTCTTCCGGGGTCTCGCAGTATAGTGGCGCGCGAGGGGAGGCGCGG 300
Qy 301 TGGCGGTACGCACTGGCCAGCGGCATTTGGCGGTGGGTGGCGGGCGGCGGGAATGG 360
Db 301 TGGCGGTACGCACTGGCCAGCGGCATTTGGCGGTGGGTGGCGGGCGGCGGGAATGG 360
Qy 361 AGCAGCGCCCGGGAGCGCCGCGACAGCCAGTGAAGGAGAGGCTGTAATGTCGAACAT 420
Db 361 AGCAGCGCCCGGGAGCGCCGCGACAGCCAGTGAAGGAGAGGCTGTAATGTCGAACAT 420
Qy 421 TGGGAATTCGCACTTAATATTTATGGCACAAATTCGTTGATGGACAATCTGATAGACAAA 480
Db 421 TGGGAATTCGCACTTAATATTTATGGCACAAATTCGTTGATGGACAATCTGATAGACAAA 480
Qy 481 ATTGCTGGGACAAAAGGCTGTCTGATATGGATAACAGGACTCAGTGGTTTCAGGGAAGAAG 540
Db 481 ATTGCTGGGACAAAAGGCTGTCTGATATGGATAACAGGACTCAGTGGTTTCAGGGAAGAAG 540
Qy 541 TACTCTTGCACTGTCAGTCTGAGTTGCAATTCGACAGAGCCACCTCAGCTATGTACT 600
Db 541 TACTCTTGCACTGTCAGTCTGAGTTGCAATTCGACAGAGCCACCTCAGCTATGTACT 600
Qy 601 TGATGGTGACAACTCAGACATGGCCCTAAATAGAGATTTAAGCTTTTAAGCGAAGACCG 660
Db 601 TGATGGTGACAACTCAGACATGGCCCTAAATAGAGATTTAAGCTTTTAAGCGAAGACCG 660

Qy 661 TGCAGAAAATATACGAAGATTGGTGAAGTGCACAAAGCTTTTGGTGTATGCTGGTGTCTCAT 720
Db 661 TGCAGAAAATATACGAAGATTGGTGAAGTGCACAAAGCTTTTGGTGTATGCTGGTGTCTCAT 720
Qy 721 ATGCATTGCTAGCTTTGATATCTCCATACAGAGAGATCGTGATGATCGCCGCTCTACT 780
Db 721 ATGCATTGCTAGCTTTGATATCTCCATACAGAGAGATCGTGATGATCGCCGCTCTACT 780
Qy 781 TCCACATTCTAACTTTTATTTGAAGTATTTATTTGCCCCCTTAAATTTTGTGAAGCTCG 840
Db 781 TCCACATTCTAACTTTTATTTGAAGTATTTATTTGCCCCCTTAAATTTTGTGAAGCTCG 840
Qy 841 TGATCTTAAAGCCCTATACAGCTTGCACGTACAGAAAGATTTAAAGTTTTCACCTGGAAT 900
Db 841 TGATCTTAAAGCCCTATACAGCTTGCACGTACAGAAAGATTTAAAGTTTTCACCTGGAAT 900
Qy 901 TGATGATCCATACGAAACCAATTAATGCTGAGATAGTAAATTAAGATGAAGATGAGGA 960
Db 901 TGATGATCCATACGAAACCAATTAATGCTGAGATAGTAAATTAAGATGAAGATGAGGA 960
Qy 961 ATGCCCTTCAACCAAGCAATGGCCAAAGCTTCTATGCTACCTTGAAGAAACGGATA 1020
Db 961 ATGCCCTTCAACCAAGCAATGGCCAAAGCTTCTATGCTACCTTGAAGAAACGGATA 1020
Qy 1021 TTTGCAAGCTTAGTATATGATTTTGAAGATTTGATCTGTTGTTGCTGCCATTACT 1080
Db 1021 TTTGCAAGCTTAGTATATGATTTTGAAGATTTGATCTGTTGTTGCTGCCATTACT 1080
Qy 1081 TGTGGACACAATTAAGATCTGTTGTTGTCACATGAATAAAGGCATCAACATGTAGGAAG 1140
Db 1081 TGTGGACACAATTAAGATCTGTTGTTGTCACATGAATAAAGGCATCAACATGTAGGAAG 1140
Qy 1141 TAACAGAAGTACGGTTTCATTTCAGAAACGGATATGGATTTCGTTTAAATAAAAAA 1200
Db 1141 TAACAGAAGTACGGTTTCATTTCAGAAACGGATATGGATTTCGTTTAAATAAAAAA 1200
Qy 1201 AAAAAA 1217
Db 1201 AAAAAA 1217

RESULT 2
AAZ50163
ID AAZ50163 standard; cDNA: 928 BP.
XX AAZ50163;
AC AAZ50163;
DT 04-MAY-2000 (first entry)
XX Wheat Adenylylsulphate kinase-1 cDNA clone.
XX Adenylylsulphate kinase: Adenosine-5'phosphosulphate kinase; APS kinase;
XX 3'-phospho-adenosine-5'phosphosulphate; PAPS; sulphate assimilation;
KW wheat; clone wrl.pk0101.e2; transgenic plant; screen; antibody; ss.
XX Triticum aestivum.
XX
FH Key Location/Qualifiers
FT CDS 3..743
FT /*tag= a
FT /product= "Wheat APS kinase-1"
FT /note= "Derived from clone wrl.pk0101.e2"
XX
XX WO200004165-A1.
XX
XX 27-JAN-2000.
XX
XX 13-JUL-1999; 99WO-US15809.
XX
XX 14-JUL-1998; 98US-0092833.
XX
XX (DUPO) DU PONT DE NEMOURS & CO E I.
XX

PI Falco SC, Allen SM, Anderson SL;
XX WPI: 2000-182430/16.
DR P-PSDB; AAY44792.
XX
PT New nucleic acid molecule and chimeric gene encoding an adenosine-5'
PT adenylylsulphate kinase, useful for altering expression of sulfate
PT assimilation protein in plants -
XX
PS Claim 3; Page 34-35; 42pp; English.
XX
XX The present sequence is a cDNA encoding wheat adenylylsulphate kinase
CC (APS kinase), also known as adenosine-5' phosphosulphate kinase. This is
CC obtained from clone wrl.pk0101.e2, derived from 7 day old light grown
CC wheat root seedlings, wrl cDNA library. APS kinase is a sulphate
CC assimilation protein, that catalyses the conversion of adenosine-5'
CC phosphosulphate (APS) to 3'-phospho-adenosine-5' phosphosulphate (PAPS).
CC This sequence is used as a probe and primer to identify, obtain and
CC synthesise sulphate assimilation proteins from other plants. It is also
CC used to produce transgenic plants, that are useful for altering the
CC expression levels of a sulphate assimilation protein. The APS kinase
CC peptides are useful for producing antibodies, that are used to screen
CC and isolate cDNA clones.
XX
SQ Sequence 928 BP; 285 A; 168 C; 250 G; 225 T; 0 other;

Query Match 29.5%; Score 359.6; DB 21; Length 928;
Best Local Similarity 65.5%; Pred. No. 1.1e-80;
Matches 558; Conservative 0; Mismatches 289; Indels 5; Gaps 2;

QY 367 GCCCGGGGAGCCCGCCACAGCCAGCTGAAGGAGAGCCTGTAATGTCGAACATTGGGAA 426
DB 77 GCCCGTCAATGGATCCCATGCAGGTATCGACAGCTTGTGACCTCAACTGTGGGAA 136
QY 427 ATCGACTAATATTTATGGCACAAATTCGTTGATTGGACAATCTCATAGACAGAAATGCT 486
DB 137 ATCGACAACCTTTCTTGGCATGACTGTCCAAATAGGTGAGTTCAGTTCAGAGGCGAAGTCT 196
QY 487 GGGACAAAAGGCTGTCTGATGATACAGGACTCAGTGGTTACAGGAAAAGTACTCT 546
DB 197 AATCAGRAGGGTTGTTGTGTGGATACAGGGTTAAGTGGTTCAGGAAAAGCACACT 256
QY 547 TCGATGTCAGTGAAGTTCGATGATGATGATGATGATGATGATGATGATGATGATGATG 606
DB 257 AGCATGCGCGTAAAGTTCGAGAGTGCAGTCCAGAGGTCATCTGACCTACATTTCTAGACGG 316
QY 607 TGACAACCTCAGACATGGCCCTAAATAGAGATTTAAGCTTTAAGGCAAGACCGTGCAGA 666
DB 317 TGACAATCAGGATGGGTTAACCAGAGCCTCTGTTCCAGCAAGGACCGTGTGA 376
QY 667 AAATATCAGAGAGTTGGTGAAGTGGCAAGCTTTTTCGTGATGCTGGTGTATATGAT 726
DB 377 AAATATCAGAGAGTAGGAGAGTAGCAAGCTGTTTCAGATGCTGGTGTATCTGAT 436
QY 727 TGTAGCTTGATATCTCCATACAGGAGAGATCGTATGATGATGATGATGATGATGATGAT 786
DB 437 TGTAGCTTGATATCCCTACAGAGTGAACGAGCGCTTGCGCAAAATTAAGTGCACAA 496
QY 787 TTCTAATCTTATGAAGTATTTATGATTTGGCCCTCAAAATTTGGAGCTGCTGATGCC 846
DB 497 TTCTAATCTCAGAGTGTGTTTGAATGTCCTCACTTTGAAGTGTGAGCTAGGATGCC 556
QY 847 TAAAGGCTTACAGCTTCGACGTACAGGAAGATTTAAGGTTTCACTGGAATTCATGA 906
DB 557 ARAAGCTTGTACAAAGCTTCGCCGTCCAGGAATAATCAAGGGTTTACTGGAATTCATGA 616
QY 907 TCATACGAACCCCAATTAATGGTGAATAGTATTAATGAATGAAGATGAGGATGCC 966
DB 617 TCCTTATGAAGCACTTCTGACTGCGAGATAGTATACAGTGCAGAACTGGTCACTGGCC 676
QY 967 TTCACCAAGCAATGCCAAGCACTTCTATGCTACCTTCAAGAAACGGATATTGCA 1026
DB 677 CAGCGCTAAATCGATGCTCATCAAGTTGTGTCATATCTTGAAGCAANTGAGTCTTACA 736

QY 1027 AGCTTAG---TATATGTATTTTGAGAGATGATCTGATCTCTGTGTCTCCATTACTTGT 1083
DB 737 GGAATAGAGACGTATGCTATGGATGAAAAACATTTCTGAAATTTGCGCAAGGGATGT 796
QY 1084 GGACACAATAAGATCTGTTGGTCCACATGAATAA--AAGGCATCAACATGTAGGAAGT 1141
DB 797 GAATATGAGGTAGTATTTATGCTCTAGAAAGAGTGTATGATGATGAGACATATATAT 856
QY 1142 AACAGAAGTACGGTTTCATTTCAGAAACGGATATGATGATTCATTCGTTTAAAAA 1201
DB 857 GACATAAGATCGAATCTGTACATCATATATAATAATGAAATGTTTTCAGCGCAAAAA 916
QY 1202 AAAAAA 1213
DB 917 AAAAAA 928

RESULT 3
AAZ50162
ID AAZ50162 standard; cDNA; 936 BP.
XX
AC AAZ50162;
XX
XX 04-MAY-2000 (first entry)
XX
XX Soybean Adenylylsulphate kinase cDNA clone.
XX
XX Adenylylsulphate kinase; Adenosine-5'phosphosulphate kinase; APS kinase;
KW 3'-Phospho-adenosine-5'phosphosulphate; PAPS; sulphate assimilation;
KW soybean; clone sdp2c.pk013.all; transgenic plant; screen; antibody; ss.
XX
OS Glycine max.
XX
FH Key Location/Qualifiers
FT CDS 1..936
FT /tag= a
FT /product= "Soybean APS kinase"
FT /note= "Derived from clone sdp2c.pk013.all"
XX
XX WO200004165-A1.
XX
XX 27-JAN-2000.
XX
XX 13-JUL-1999; 99WO-US15809.
XX
XX 14-JUL-1998; 98US-0092833.
XX (DUPO) DU PONT DE NEMOURS & CO E I.
XX
XX Falco SC, Allen SM, Anderson SL;
XX
XX WPI: 2000-182430/16.
XX P-PSDB; AAY44791.
XX
XX New nucleic acid molecule and chimeric gene encoding an adenosine-5'
PT phosphosulphate kinase, useful for altering expression of sulfate
PT assimilation protein in plants -
XX
XX Claim 3; Page 33; 42pp; English.
XX
XX The present sequence is a cDNA encoding soybean adenylylsulphate kinase
CC (APS kinase), also known as adenosine-5' phosphosulphate kinase. This is
CC obtained from clone sdp2c.pk013.all, derived from soybean developing
CC pods, sdp2c cDNA library. APS kinase is a sulphate assimilation
CC protein, that catalyses the conversion of adenosine-5' phosphosulphate
CC (APS) to 3'-phospho-adenosine-5' phosphosulphate (PAPS). This sequence
CC is used as a probe and primer to identify, obtain and synthesise sulphate
CC assimilation proteins from other plants. It is also used to produce
CC transgenic plants, that are useful for altering the expression levels of
CC a sulphate assimilation protein. The APS kinase peptides are useful for
CC producing antibodies, that are used to screen and isolate cDNA clones.
XX

PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161992.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.

Query Match

Best Local Similarity

Matches 435; Conservative

26.3%; Score 320.4; DB 21; Length 627;

Pred. No. 7e-71;

Mismatches 191; Indels 0; Gaps 0

Qy	410	ATGTCGAACATTGGGAAATCGCACTAATAATTTATGCGCACAAATTCGTGATGCACAATTCGTGATGGACAATCT	469
Db	1	ATGTCGACAGTGGGAAATTC AACGAAACATAATTTTGCGAAGAATCCCCCATTTGGGAAAAC T	60
Qy	470	GATACGACAAANTCTGGGCAAAAGGCTGCTGATGCATACAGACTCAGTGGT	529
Db	61	GAAGGCGAAGTGTGTTAAATCAGANGGGTGTGTGTTGGATACAGGGCTCAGTGGC	120
Qy	530	TCAGGGAAGAATACTCTTGCAATGTCGACTGAGTCGTGAGTTGCAATTCAGAGGCCA C CTC	589
Db	121	TCAGGGAAGACAGTTAGCTTGCCTTAGTAGAGAGCTGAAACACC GGGAAGCTA	180
Qy	590	ACGTATGTACTTGTGTCGACAACTCAGACCTGCGCTAAATAGAGATTTAAGCTTTTAAG	649
Db	181	TCAATATCTTGATGGTGACNACTTCGTCATGGTTGCAACAAGATCTTGTTCCAAG	240
Qy	650	GCAGAAGACCGTCGAGAAATATACGAAGAGTTGGTGAAGTGGCAAGCTTTTTCGTGAT	709
Db	241	GCAGAGGATAGTGGAAATATACGAGGGTGGGAAGTAGCCAACATCTTTTCGCGAT	300
Qy	710	GCTGTGTGCATATGCATTCTGTAGCTTGATATCTCCATACAGGAGAGTCGTGATGCATGC	769
Db	301	GCTGTGTTAATCTGTATTGCCAGCCTCATATCCCCTATAGAAAGACC GTGACCCCTGC	360
Qy	770	CSTGCTCTACTTCCACATTTCTAACTTTTATTGAAGTATTTATTGATTTGCCCTAAAAAT	829
Db	361	AGGGAATGATACAGAAATTCATCTTTTATTAGGTGTTTCATGACATGTCTCTGCAATTG	420
Qy	830	CTGTAGCTGTGATCCTTAAGGCTATACAAGCTTGTACAGCTCAGAGAAAGATTAAGGT	889
Db	421	TGTGAAGCAGGAGCCCTTAAGGCTATACAGCTTGCACGTGCAGGAAGATCAAAGT	480
Qy	890	TTCACTTGGAAATGATGATCCATPACGAACCACCAATTTAATGTGTGAGTATTAAGATG	949

PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 26-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.

PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144332.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148171.
PR 12-AUG-1999; 99US-0148319.
PR 13-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149920.
PR 23-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.

PR	07-OCT-1999;	99US-0158029.	
PR	08-OCT-1999;	99US-0158232.	
PR	12-OCT-1999;	99US-0158369.	
PR	13-OCT-1999;	99US-0159293.	
PR	13-OCT-1999;	99US-0159294.	
PR	13-OCT-1999;	99US-0159295.	
PR	14-OCT-1999;	99US-0159329.	
PR	14-OCT-1999;	99US-0159330.	
PR	14-OCT-1999;	99US-0159331.	
PR	14-OCT-1999;	99US-0159637.	
PR	14-OCT-1999;	99US-0159638.	
PR	18-OCT-1999;	99US-0159584.	
PR	21-OCT-1999;	99US-0160741.	
PR	21-OCT-1999;	99US-0160767.	
PR	21-OCT-1999;	99US-0160768.	
PR	21-OCT-1999;	99US-0160770.	
PR	21-OCT-1999;	99US-0160814.	
PR	21-OCT-1999;	99US-0160815.	
PR	22-OCT-1999;	99US-0160980.	
PR	22-OCT-1999;	99US-0160981.	
PR	22-OCT-1999;	99US-0160989.	
PR	22-OCT-1999;	99US-0161404.	
PR	23-OCT-1999;	99US-0161405.	
PR	26-OCT-1999;	99US-0161406.	
PR	26-OCT-1999;	99US-0161359.	
PR	26-OCT-1999;	99US-0161360.	
PR	26-OCT-1999;	99US-0161361.	
PR	28-OCT-1999;	99US-0161920.	
PR	28-OCT-1999;	99US-0161992.	
PR	28-OCT-1999;	99US-0161993.	
PR	29-OCT-1999;	99US-0162142.	
Query Match			
Best Local Similarity 24.4%; Score 296.4; DB 21; Length 1068;			
Mismatches 430; Conservative 68.0%; Pred. No. 9.7e-65;			
Matches 430; Indels 0; Gaps 1;			
QY	411	TGTCGAACATGGGAATGACATTAATATTTATGSCACATTCCTTGATTGGACAACTCG	470
DB	308	TGTCACGGTCGGAATCGCAATATTAAGTGCATGATGTCTGTGTGAGAAAGTTG	367
QY	471	ATAGACAGAAATGCTGGGACAAAAGGCTGTGCTGATGATGAATACAGGACTCAGTGGTT	530
DB	368	ATAGACAGAGATTCTTGATGACAGAAAGGATGTGATTTGGTCAACGGCTCTTAGTGGTT	427
QY	531	CAGGAAAGTACTCTTGGCATGTCAGTGTGATGCTGATGCTGATGCTGACAGGCCACCTCA	590
DB	428	CAGGAAAGTACTCTTGGCTTGTGCTTTGAATCAGATGTTGATCAAAAGGGGAACCTTT	487
QY	591	CGTATGTTACTGTGTCAGACCTCAGACATGGCCTAAATAGAGATTTAAGCTTTAAGG	650
DB	488	GTATATATCTTGTGTTGATGTAATGTTAGGCATGGCTTAAACCGTGATCTTAGCTTTAAG	547
QY	651	CAGAAGACCGTGCAGAAAATATACGAAGTTGGTGAAGTGGCAAAAGCTTTTGTCTGATG	710
DB	548	CTGAGGATCGTCAGAGAAATATCGTAGAGTTGGAGAGTTGCTAAGCTTTTGGCGATG	607
QY	711	CTGTGTCTATGATGCTAGCTTGATCTCCATCAGAGAGATCGTGTGATGCTGCC	770
DB	608	CTGGAATATCTGATGCGAGTTTGATATCTCTTATAGAACAGATAGGACGCTTCTC	667
QY	771	GTGCTCTACTTCCACATCTTAACCTTATTTGAAGTATTTATGATTTGGCCCTAAAAATTT	830
DB	668	GAAGTTTCTCCCGAGGGAGATTTTGTGAGGTGTTTCATGGATGTACCGCTTAGTGT	727
QY	831	GTGAAGCTGTGATCTAAAGGCTTATACAGCTTGCACGTACAGGAAGATTAAGTTT	890
DB	728	CGGAGGCGAGGATCCAAAGGGTCTTTAAAGCTTGCTGTCGAGGAAGATCAAAAGTT	787
QY	891	TCACCTGGAATTTGATGATCCATCAGCAACCAATTAATGTTGATAGATAGTAATTAAGATGA	950
DB	788	TTACCGGGATGATGACCTTAGGACCCACCATTTGAACCTCGAG-----ATTCTCTAG	841
QY	951	AGATGAGGAATGCCCTTACCCAAAGCAATGGCCAAAGTTCCTATGCTACCTTGAAG	1010

DB	842	GACGTGAAGGAGAACTTCTCTATCGAAATGCGGAAAGTCTCGGATACTTAGATA	901
QY	1011	AAACGGATATTTGCAAGCTTAGTATATGTAT	1042
DB	902	ACAAGGGTTATCTTCAAGCATACATACTTCT	933
RESULT 7			
ID	AAC36588		
XX	AAC36588 standard; DNA; 1175 BP.		
AC	AAC36588;		
XX			
DT	17-OCT-2000 (first entry)		
XX			
DE	Arabidopsis thaliana DNA fragment SEQ ID NO: 14353.		
XX			
KW	Hybridisation assay; genetic mapping; gene expression control;		
KW	protein identification; signal transduction pathway;		
KW	metabolic pathway; promoter; termination sequence; ss.		
XX			
OS	Arabidopsis thaliana.		
XX			
PN	EF1033405-A2.		
XX			
PD	06-SEP-2000.		
PF			
XX	25-FEB-2000; 2000EP-0301439.		
PR	25-FEB-1999;	99US-0121825.	
PR	05-MAR-1999;	99US-0123180.	
PR	09-MAR-1999;	99US-0123548.	
PR	23-MAR-1999;	99US-0125788.	
PR	25-MAR-1999;	99US-0126264.	
PR	29-MAR-1999;	99US-0126785.	
PR	01-APR-1999;	99US-0127462.	
PR	06-APR-1999;	99US-0128234.	
PR	08-APR-1999;	99US-0128714.	
PR	16-APR-1999;	99US-0129845.	
PR	19-APR-1999;	99US-0130077.	
PR	21-APR-1999;	99US-0130449.	
PR	23-APR-1999;	99US-0130510.	
PR	28-APR-1999;	99US-0130891.	
PR	30-APR-1999;	99US-0131449.	
PR	30-APR-1999;	99US-0132048.	
PR	30-APR-1999;	99US-0132407.	
PR	04-MAY-1999;	99US-0132484.	
PR	05-MAY-1999;	99US-0132485.	
PR	06-MAY-1999;	99US-0132486.	
PR	07-MAY-1999;	99US-0132487.	
PR	11-MAY-1999;	99US-0132863.	
PR	14-MAY-1999;	99US-0134256.	
PR	14-MAY-1999;	99US-0134218.	
PR	14-MAY-1999;	99US-0134219.	
PR	14-MAY-1999;	99US-0134221.	
PR	14-MAY-1999;	99US-0134370.	
PR	18-MAY-1999;	99US-0134768.	
PR	19-MAY-1999;	99US-0134941.	
PR	20-MAY-1999;	99US-0135124.	
PR	21-MAY-1999;	99US-0135353.	
PR	24-MAY-1999;	99US-0135629.	
PR	25-MAY-1999;	99US-0136021.	
PR	27-MAY-1999;	99US-0136392.	
PR	28-MAY-1999;	99US-0136782.	
PR	01-JUN-1999;	99US-0137222.	
PR	03-JUN-1999;	99US-0137528.	
PR	04-JUN-1999;	99US-0137502.	
PR	07-JUN-1999;	99US-0137724.	
PR	08-JUN-1999;	99US-0138094.	
PR	10-JUN-1999;	99US-0138540.	
PR	10-JUN-1999;	99US-0138847.	
PR	14-JUN-1999;	99US-0139119.	

[illegible]

PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134321.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 22-JUN-1999; 99US-0139817.
PR 23-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0143055.
PR 06-JUL-1999; 99US-0143390.
PR 08-JUL-1999; 99US-0144803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149721.
PR 20-AUG-1999; 99US-0149923.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.

XX		WO200008187-A2.	
PX		17-FEB-2000.	
PD			
PP		04-AUG-1999; 99WO-EPO6562.	
PF			
PR		04-AUG-1998; 98EP-0202634.	
PS	(VLA--)	VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.	
PA		Lee JH, Verbruggen N;	
PI		WPI: 2000-205726/18.	
DR		p-PADB; AAY77957.	
DD			
XX		Isolation of polynucleic acids useful for producing transgenic plant isolating genes involved in tolerance to environmental stress - Claim 4; Page 218-220; 312pp; English. The invention relates to isolation of coding sequences and/or genes involved in tolerance to environmental stress in plants. The sequence (AA298305-298365) are useful for producing a transgenic plant having enhanced tolerance or resistance to environmental stress conditions such as anaerobic, flooding, cold, dehydratation, drought, heat stress or salinity. This is useful for producing improved yield, growth, development and productivity under environmental stress conditions, also also provides growth of crops in areas where they cannot grow without the induced osmotolerance. Sequences AA298305-365 represent polynucleotide sequences from A. thaliana that are involved in environmental stress tolerance. Sequence 917 BP; 244 A; 220 C; 229 G; 224 T; 0 other; Query Match 21.7%; Score 263.6; DB 21; Length 917; Best Local Similarity 65.5%; Pred. No. 1.7e-56; Matches 402; Conservative 0; Mismatches 209; Indels 3; Gaps	
QY	422	GCGAATCAGCTTAATAATTATTTGGACCAATTCGTGTGATGGACAANTCTCATAGACAAA	481
Dd	293	GAGAAGCAGAGAACATCGTGTGCACGAGATTCGATATCGAGTGCGACAGACAA	352
QY	482	TTCGTGGGCAAAGAGGCTGTCGTATGVATAACAGACTCAGTGGTTCCAGGAAAAAGT	541
Dd	353	CCTCTTCAACAAGGGTGTGCTGTGGATCACTGTCTCAGTGGTTCAGSNAAGC	412
QY	542	ACTCTTCGATGTCAGCTGFCGTGGTAGTTCGATTCGACGCCACCCTCACGTATGTACTT	601
Dd	413	ACTGTGCTTGTGCACGAAGTAAGACATGTTTTGAAGAGGCAAACTTACTTTACACACTC	472
QY	602	GATGSTGACAACTCAGACATGGCTTAATAGAGATTTAAGCTTTAAGSCAGAGACGCGT	661
Dd	473	GACGGGACAAATGTCGTCACGGCTTTAACCGTGACCTCACTTCCAAGCAGACGACCGC	532
QY	662	GCAGAAATATACGAGAGTGTGTAAGTGCGCAAGACTTTTTCGTGATCGTGTCATATA	721
Dd	533	ACCGAAACATATAGAAATTTGTCAGTGGCTAGTTGTTGCTGCAGCTCGAGTCAAT	592
QY	722	TGCATFTGCTAGCTTGATATCTCCATACAGAGAGATCTGTGATGATCGCGTCTCTACTT	781
Dd	593	TGATACCAAGTTTGATTTCTCCGTACCGGAGACAGACGCGGTCGCGGTCCTTGTA	652
QY	782	CCACATTTCTAACTTTATTGAAGTATTTATTGATTTGCCCTTAAAAATTTGTGAAGCTCGT	841
Dd	653	CCTGACGGGATTTGCTCGAGSTCTTCTATGAGCTTCTCTTCAATGTGTGCGAGTCGAGA	712
QY	842	GATCCTTAAGGCCATACAGCTTGCAGCTACAGAAAGATTAANGGTTTCACTGGRATT	901
Dd	713	GATCAAGAAGGGGTGTGACAAAGCTTGACAGCTPGCAGGCAAAATCAAGGCTTCACTGGRAIC	772
QY	902	GATGATCCATACGAACCCCACCAATTAATGTGTGAGATAGTAATTA---AGATAAAGATGAG	958

```
Db 773 GACGACCTTACGAGGCACCACTGATTCGCGAGGTAGTGCTGAAACACACAGCAGCAC 832
Qy 959 GAATGCCCTTACCAACCAAGCAATGGCCCAAGCAAGTTCTATGCTTACCTTGAAGAAAACGGA 1018
Db 833 GATCGTGTTCGCACCTCAGATGGCTGAGAACATCATCTCTTACCTGCAAAACAAGGT 892
Qy 1019 TATTTGCAAGCTTA 1032
Db 893 TATCTTGAGGCTA 906

RESULT 12
AAC42294
ID AAC42294 standard; DNA; 948 BP.
XX
AC AAC42294;
XX
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 35008.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123348.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-01355629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 06-JUL-1999; 99US-0142803.
PR 08-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
```


Query Match	Score	263 6;	DB 21;	Length	948;
Best Local Similarity	65.5%				
Matches 402;	Conservative	0;	Mismatches 209;	Indels	3;
Gaps	1;				
<p> 99US-0147935. 99US-0148171. 99US-0148319. 99US-0148341. 99US-0148565. 99US-0148684. 99US-0149368. 99US-0149175. 99US-0149426. 99US-0149722. 99US-0149723. 99US-0149929. 99US-0149902. 99US-0149930. 99US-0150566. 99US-0150884. 99US-0151065. 99US-0151066. 99US-0151080. 99US-0151303. 99US-0151438. 99US-0151930. 99US-0152363. 99US-0153070. 99US-0153758. 99US-0154018. 99US-0154039. 99US-0154779. 99US-0155139. 99US-0155486. 99US-0155659. 99US-0156458. 99US-0156596. 99US-0157117. 99US-0157753. 99US-0157865. 99US-0158029. 99US-0158232. 99US-0158369. 99US-0159293. 99US-0159294. 99US-0159295. 99US-0159329. 99US-0159330. 99US-0159331. 99US-0159637. 99US-0159638. 99US-0159584. 99US-0160741. 99US-0160767. 99US-0160768. 99US-0160770. 99US-0160814. 99US-0160815. 99US-0160980. 99US-0160981. 99US-0160989. 99US-0161404. 99US-0161405. 99US-0161406. 99US-0161359. 99US-0161360. 99US-0161361. 99US-0161920. 99US-0161992. 99US-0161993. 99US-0162142. </p>					

PR	06-APR-1999;	99US-0128234.	PR	19-JUL-1999;	99US-0144332.
PR	08-APR-1999;	99US-0128714.	PR	19-JUL-1999;	99US-0144333.
PR	16-APR-1999;	99US-0128845.	PR	19-JUL-1999;	99US-0144334.
PR	19-APR-1999;	99US-0130077.	PR	20-JUL-1999;	99US-0144335.
PR	21-APR-1999;	99US-0130449.	PR	20-JUL-1999;	99US-0144632.
PR	23-APR-1999;	99US-0130510.	PR	20-JUL-1999;	99US-0144884.
PR	23-APR-1999;	99US-0130891.	PR	21-JUL-1999;	99US-0144814.
PR	28-APR-1999;	99US-0131449.	PR	21-JUL-1999;	99US-0145086.
PR	30-APR-1999;	99US-0132048.	PR	21-JUL-1999;	99US-0145088.
PR	04-MAY-1999;	99US-0132407.	PR	21-JUL-1999;	99US-0145088.
PR	05-MAY-1999;	99US-0132484.	PR	22-JUL-1999;	99US-0145087.
PR	06-MAY-1999;	99US-0132486.	PR	22-JUL-1999;	99US-0145089.
PR	06-MAY-1999;	99US-0132487.	PR	22-JUL-1999;	99US-0145192.
PR	07-MAY-1999;	99US-0132863.	PR	23-JUL-1999;	99US-0145145.
PR	11-MAY-1999;	99US-0134256.	PR	23-JUL-1999;	99US-0145218.
PR	14-MAY-1999;	99US-0134218.	PR	23-JUL-1999;	99US-0145224.
PR	14-MAY-1999;	99US-0134219.	PR	26-JUL-1999;	99US-0145276.
PR	14-MAY-1999;	99US-0134221.	PR	27-JUL-1999;	99US-0145913.
PR	14-MAY-1999;	99US-0134370.	PR	27-JUL-1999;	99US-0145918.
PR	18-MAY-1999;	99US-0134768.	PR	27-JUL-1999;	99US-0145919.
PR	19-MAY-1999;	99US-0134941.	PR	28-JUL-1999;	99US-0145951.
PR	20-MAY-1999;	99US-0135124.	PR	02-AUG-1999;	99US-0146386.
PR	21-MAY-1999;	99US-0135353.	PR	02-AUG-1999;	99US-0146388.
PR	21-MAY-1999;	99US-0135629.	PR	02-AUG-1999;	99US-0146389.
PR	25-MAY-1999;	99US-0136021.	PR	03-AUG-1999;	99US-0147038.
PR	27-MAY-1999;	99US-0136392.	PR	04-AUG-1999;	99US-0147204.
PR	28-MAY-1999;	99US-0136782.	PR	04-AUG-1999;	99US-0147302.
PR	01-JUN-1999;	99US-0137222.	PR	05-AUG-1999;	99US-0147192.
PR	03-JUN-1999;	99US-0137528.	PR	05-AUG-1999;	99US-0147260.
PR	04-JUN-1999;	99US-0137502.	PR	06-AUG-1999;	99US-0147303.
PR	07-JUN-1999;	99US-0137724.	PR	06-AUG-1999;	99US-0147416.
PR	08-JUN-1999;	99US-0138094.	PR	09-AUG-1999;	99US-0147493.
PR	10-JUN-1999;	99US-0138540.	PR	09-AUG-1999;	99US-0147935.
PR	10-JUN-1999;	99US-0138847.	PR	10-AUG-1999;	99US-0148171.
PR	14-JUN-1999;	99US-0139119.	PR	11-AUG-1999;	99US-0148319.
PR	16-JUN-1999;	99US-0139452.	PR	12-AUG-1999;	99US-0148341.
PR	16-JUN-1999;	99US-0139453.	PR	13-AUG-1999;	99US-0148565.
PR	17-JUN-1999;	99US-0139492.	PR	13-AUG-1999;	99US-0148684.
PR	18-JUN-1999;	99US-0139454.	PR	16-AUG-1999;	99US-0149368.
PR	18-JUN-1999;	99US-0139455.	PR	17-AUG-1999;	99US-0149175.
PR	18-JUN-1999;	99US-0139456.	PR	18-AUG-1999;	99US-0149426.
PR	18-JUN-1999;	99US-0139457.	PR	20-AUG-1999;	99US-0149722.
PR	18-JUN-1999;	99US-0139458.	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	99US-0139459.	PR	20-AUG-1999;	99US-0149929.
PR	18-JUN-1999;	99US-0139460.	PR	23-AUG-1999;	99US-0149902.
PR	18-JUN-1999;	99US-0139461.	PR	23-AUG-1999;	99US-0149930.
PR	18-JUN-1999;	99US-0139462.	PR	25-AUG-1999;	99US-0150566.
PR	18-JUN-1999;	99US-0139463.	PR	26-AUG-1999;	99US-0150884.
PR	18-JUN-1999;	99US-0139750.	PR	27-AUG-1999;	99US-0151065.
PR	18-JUN-1999;	99US-0139763.	PR	27-AUG-1999;	99US-0151066.
PR	21-JUN-1999;	99US-0139817.	PR	27-AUG-1999;	99US-0151080.
PR	22-JUN-1999;	99US-0139899.	PR	30-AUG-1999;	99US-0151303.
PR	23-JUN-1999;	99US-0140353.	PR	31-AUG-1999;	99US-0151438.
PR	23-JUN-1999;	99US-0140354.	PR	01-SEP-1999;	99US-0151930.
PR	24-JUN-1999;	99US-0140695.	PR	07-SEP-1999;	99US-0152363.
PR	24-JUN-1999;	99US-0140823.	PR	10-SEP-1999;	99US-0153070.
PR	29-JUN-1999;	99US-0140991.	PR	13-SEP-1999;	99US-0153758.
PR	30-JUN-1999;	99US-0141287.	PR	15-SEP-1999;	99US-0154018.
PR	01-JUL-1999;	99US-0141842.	PR	16-SEP-1999;	99US-0154039.
PR	01-JUL-1999;	99US-0142154.	PR	20-SEP-1999;	99US-0154779.
PR	02-JUL-1999;	99US-0142055.	PR	22-SEP-1999;	99US-0155139.
PR	06-JUL-1999;	99US-0142390.	PR	23-SEP-1999;	99US-0155348.
PR	08-JUL-1999;	99US-0142803.	PR	24-SEP-1999;	99US-0155659.
PR	09-JUL-1999;	99US-0142920.	PR	28-SEP-1999;	99US-0156458.
PR	12-JUL-1999;	99US-0142977.	PR	29-SEP-1999;	99US-0156596.
PR	13-JUL-1999;	99US-0143542.	PR	04-OCT-1999;	99US-0157117.
PR	14-JUL-1999;	99US-0143624.	PR	05-OCT-1999;	99US-0157753.
PR	15-JUL-1999;	99US-0144005.	PR	06-OCT-1999;	99US-0157865.
PR	16-JUL-1999;	99US-0144085.	PR	07-OCT-1999;	99US-0158029.
PR	16-JUL-1999;	99US-0144086.	PR	08-OCT-1999;	99US-0158232.
PR	19-JUL-1999;	99US-0144325.	PR	12-OCT-1999;	99US-0158369.
PR	19-JUL-1999;	99US-0144331.	PR	13-OCT-1999;	99US-0159293.

PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 26-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 18.5%; Score 225; DB 21; Length 714;

Best Local Similarity 61.2%; Pred. No. 8e-47;

Matches 436; Conservative 0; Mismatches 190; Indels 87; Gaps 1;

QY 410 ATGTGACACATTTGGAAATCGACTAATATTTATGGCACAAATGCTTGATGGACAATCT 469
DB 1 ATGTGACAGTGGAAATTCACGACACATATTTTGGCAAGAAATCCCCCAATGGGAAACT 60
QY 470 GATAGACAGAAATTTGGGACAAAAGCGTGTGCTGATGGATACAGAGCTCAGTGGT 529
DB 61 GAAAGCCAGAAAGTTGTTAAATCAGAAAGGTTGTGTGGTGTGGATCAGAGGCTCAGTGGC 120
QY 530 TCAGGGAAGAACTACTCTGACATGAGTGTGATGAGTTGATTCAGAGGCCACCTC 589
DB 121 TCAGGGAAGAAAGCTGTGCTGCTTGTAGTAGAGCTGACACACCGGGGAAAGCTA 180
QY 590 ACGTATGACTGTGATGACACACCTCAGACATGACGCTTAATAGATGATTAAGCTTTAAG 649
DB 181 TCATATATTTCTGTATGTCGACAACTCTTCGTCATGCTGTTTGAACAAAGATCTTGGTTTCAAG 240
QY 650 GCAGGAAGACCTGCGAGAAATATACGAAGAT----- 681
DB 241 GCAGAGGATAGTAGGGGAAATATACGAGGTCGTTGATGAAAGCTGATTCGAGAGAA 300
QY 682 -----T 682
DB 301 CTCCTAATAACAATGCAAAATCCTTTTAAAGAGTGTCTCATATGATGATAGTATTACAT 360
QY 683 GGTGAAGTGGCAAGCTTTTGTGATGCTGTGTCATATGCTATGCTAGCTTGATATCT 742
DB 361 GGAGAGTAGCAAACTCTTGGGATGCTGTTTAACTGATTCGAGGCTCATATCC 420
QY 743 CCATACAGAGAGATCGTATGATGCGCGTCTCTACTTCCACATCTTAACCTTTATGAA 802
DB 421 CCGTATAGAAAGACCGTCGACGCTCGAGGAAATGATACAGAATTCATCTTTATTGAG 480
QY 803 GTATTTATTGATTTGGCCCTTAAATTTGTGAAGCTGTGATCTTAAAGSCCTATACAG 862
DB 481 GTTTTCATGACATGCTCTGCAATTTGTGAGCAAGGACCCCTAAGCCCTATACAG 540
QY 863 CTTGCACCTACAGGAAGATTAAGGTTTCACTGGAATGATGATCCATACGAACCA 922
DB 541 CTTGCACCTGAGGAAGATCAAGGTTTTCACAGGAATAGATCCGATGATGCTCCC 600
QY 923 ATTATGCTGAGATAGTAAATTAGATGAAAGATGAGGAATGCCCTTACCACCAAGCAATG 982

DB 601 TTGAAGTGTGAGATAGAGTTGAAAGAGAGAGGAGAGTGTCTTCCCTGTAGTATG 560
QY 983 GCCAAGCAAGTTCTATGCTACCTTTGAGAAAAAGGATATTTGCAAGCTTAGTA 1035
DB 661 GCTGAGGAAGTGTCTCTTATCTAGAAAGACAAAGGTTTCTTCAAAACGAGTA 713
RESULT 14
AAC51315
ID AAC51315 standard; DNA; 483 BP.
XX
AC AAC51315;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 68060.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126284.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.

```

PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139464.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 22-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0143920.
PR 12-JUL-1999; 99US-0143977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.

```

```

PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160981.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

Query Match 16.3%; Score 198.2; DB 21; Length 483;
 Best Local Similarity 65.8%; Pred. No. 3.8e-40;
 Matches 318; Conservative 0; Mismatches 138; Indels 27; Gaps 1;

QY 410 ATGTGCAACATTGGGAATCGACTAATTTTATGGCAATTCCTTTGTTGGACAATCT 469
 DB 1 ATGTGCAACATTGGGAATTCACGACATATTTTGGCAAGATCCCCCATTTGGGAATCT 60

```
QY 470 GATAGACAGAAATTTGCTGGGACAAAGGCTGTGCTATGGATTAACAGAGCTCAGTGGT 529
DB 61 GAAGGAGAGAGTTTAAATCAGAAAGGTTTGTGCTGATCACAGGGCTCAGTGGC 120
QY 530 TCAGGAAAGTACTCTTGCATGTGCTGAGTGTGAGTTGCATTCAGAGGCCACCTC 589
DB 121 TCAGGAAAGCAGCTTGTGCTGCTAGTAGAGAGCTGAACAACCGGGGAAGCTA 180
QY 590 AGCTATGATGTGATGTCAGACCTCAGACATGGCCCTAAATAGAGATTAAAGCTTTAAG 649
DB 181 TCATATATCTTGTATGGTGACAACTCTGCTCATGTTTGAACAAAGATCTTGGTTTCAAG 240
QY 650 GCAGAAAGACCGTGCAGAAAATATACGAAGAGTTGCTGAAGTGGCAAGCTTTTTCCTGAT 709
DB 241 GCAGAGGATAGAGTGGAAATATACGAGGGTCGGAGAAGTACCCAACTCTTTCGGAT 300
QY 710 GCTGTGTCATATGCTAGCTGATGATATCTCCATACAGGAGAGATCGTATCATGTC 769
DB 301 GCTGTTTAACTGTATGCGAGCTCATATCCCGTATAGAAAAGACCGTGACGCTGC 360
QY 770 CGTGCTCTACTTCCACATTTCACTTTATTGA-----A 802
DB 361 AGGGAATGATACAGAACTTCATCTTTATTGAGGCAAGCTATTCAAACCGCTTTACA 420
QY 803 GTATTTATGTTGCCCCCTAAATTTGTGAAGCTGCTGATCCTAAAGGCTATACAAG 862
DB 421 GTTTTCATGAACATGTCTGCAATTTGTGAAGCAAGGGACCCCTAAAGGCTATACAAG 480
QY 863 CTT 865
DB 481 CTT 483

RESULT 15
AAZ50161
ID AAZ50161 standard; cDNA; 431 BP.
XX
AC AAZ50161;
XX
DT 04-MAY-2000 (first entry)
XX
DE Rice Adenylsulphate kinase cDNA clone.
XX
KW Adenylsulphate kinase; Adenosine-5'-phosphosulphate kinase; APS kinase;
KW 3'-Phospho-adenosine-5'-phosphosulphate; PAPS; sulphate assimilation;
KW rice; clone r10n.pk112.011; transgenic plant; screen; antibody; ss.
XX
OS Oryza sativa.
XX
FH Key Location/Qualifiers
FT mat_peptide 55..408
FT /*tag= a
FT /product= "Rice APS kinase"
FT /note= "Derived from clone r10n.pk112.011"
XX
PN WO200004165-A1.
XX
PD 27-JAN-2000.
XX
PF 13-JUL-1999; 99WO-0515809.
XX
PR 14-JUL-1998; 98US-0092833.
XX
PA (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI Falco SC, Allen SM, Anderson SL;
XX
DR WPI; 2000-182430/16.
DR P-PSDB; AAY44790.
XX
PT New nucleic acid molecule and chimeric gene encoding an adenosine-5'
PT phosphosulphate kinase, useful for altering expression of sulfate
PT assimilation protein in plants -
```

```
XX Claim 3; Page 32; 42pp; English.
PS
XX The present sequence is a cDNA encoding rice adenylsulphate kinase
CC (APS kinase), also known as adenosine-5' phosphosulphate kinase. This is
CC obtained from clone r10n.pk112.011, derived from rice 15 day leaf, r10n
CC cDNA library. APS kinase is a sulphate assimilation protein, that
CC catalyses the conversion of adenosine-5' phosphosulphate (APS) to
CC 3'-phospho-adenosine-5' phosphosulphate (PAPS). This sequence is used as
CC probes and primers to identify, obtain and synthesise sulphate
CC assimilation proteins from other plants. It is also used to produce
CC transgenic plants, that are useful for altering the expression levels of
CC a sulphate assimilation protein. The APS kinase peptides are useful for
CC producing antibodies, that are used to screen and isolate cDNA clones.
XX
SQ Sequence 431 BP; 126 A; 78 C; 112 G; 112 T; 3 other;
```

```
Query Match 16.2%; Score 197.2; DB 21; Length 431;
Best Local Similarity 72.9%; Pred. No. 6.5e-40;
Matches 253; Conservative 0; Mismatches 94; Indels 0; Gaps 0;
QY 424 GAAATCGACTAATATTTTATGGCACAATTCGTTGATTGGACAACTCTGATAGACAGAAAT 483
DB 66 GAAGGCGTCCAATATCTTCTGGCATGATTGTCAGTTGGCCAGGCTGATCGGCAAGCT 125
QY 484 GCTGGGACAAAAGGCTGTGCTATGGATAACAGGACTCAGTGTTCAGGGAAAAAGTAC 543
DB 126 ACTGAAGCAGAAAAGGTTGCGTTGTTGGATCACAGGACTTAGTGGTTTCAGGTTAAAGTAC 185
QY 544 TCTTGCATGTCAC'TGAGTTCGAGTTGCAATTCAGAGGCCACCTCAGTATGTACTTGA 603
DB 186 CTGGGATGCACATTAGATCGAGAGCTCCATACAAGAGGGAAGCTTCTTATGTTCTTGA 245
QY 604 TGCTGACAACTCAGACATGGCCCTAAATAGAGATTTAAAGTCGAGCAAGACCGCTGC 663
DB 246 TGGTGAATAATTAAGACATGGTTTGAACAAGGATCTTGGCTTTAAGCGGGAAGACCGTGC 305
QY 664 AGAAAATATACGAAGAGTTGTTGGAAGTGGCAAGCTTTTGTGCTGATGCTGTGTCATATG 723
DB 306 TGAATATATACCAAGTTGGTGAAGTAGCAAGCTATTTCNCAGATGCAAGCCTAGTATG 365
QY 724 CATTTAGCTTGTATCTCCATACAGGAGAGATCGTGTATGATGCC 770
DB 366 CATTGCAAGTTTCAAATCTCCCTATAAGAGAGAACGCTGAGTCTCTGGC 412
```

Search completed: November 1, 2002, 23:43:31
Job time : 239 secs

GenCore version 5.1.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2002, 23:39:32 ; Search time 53 Seconds
(without alignments)
5640.301 Million cell updates/sec

Title: US-09-720-384A-3
Perfect score: 1217
Sequence: 1 ggcctcggtttcatttcatca.....aaaaaaaaaaaaaaaaaaaaa 1217

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	156.4	12.9	1160	US-09-153-310-41	Sequence 41, Appl
2	139	11.4	2506	US-08-879-561-4	Sequence 4, Appl
3	47.8	3.9	7218	US-08-232-453-14	Sequence 14, Appl
4	45.6	3.7	1323	US-09-413-452-3	Sequence 3, Appl
5	43.6	3.7	1323	US-09-413-668-3	Sequence 3, Appl
6	45	3.7	318	US-09-165-264-12	Sequence 12, Appl
7	43.4	3.6	2447	US-09-014-969-14	Sequence 14, Appl
8	42.8	3.5	12001	US-08-458-568A-11	Sequence 11, Appl
9	42.4	3.5	320	US-09-165-264-11	Sequence 11, Appl
10	42.4	3.5	320	US-09-165-264-14	Sequence 14, Appl
11	42.4	3.5	1481	US-08-474-379C-87	Sequence 87, Appl
12	42.4	3.5	3131	US-07-688-352C-21	Sequence 21, Appl
13	42.4	3.5	3131	US-09-146-249A-21	Sequence 21, Appl
14	42.4	3.5	3131	US-08-206-188B-21	Sequence 21, Appl
15	42.4	3.5	3131	PCT-US91-02714-20	Sequence 20, Appl
16	42.2	3.5	319	US-09-165-264-8	Sequence 8, Appl
17	41.4	3.4	320	US-09-165-264-7	Sequence 7, Appl
18	41.4	3.4	7218	US-08-232-453-14	Sequence 14, Appl
19	41.4	3.4	30001	US-08-125-468-1	Sequence 1, Appl
20	41.4	3.4	30001	US-08-474-933-1	Sequence 1, Appl
21	40.8	3.4	320	US-09-165-264-13	Sequence 13, Appl
22	40.6	3.3	1340	US-08-977-001-2	Sequence 2, Appl
23	40.6	3.3	4134	US-09-162-021B-1	Sequence 1, Appl
24	40.4	3.3	152331	US-09-128-155-16	Sequence 16, Appl
25	40.2	3.3	1420	US-08-909-965C-3	Sequence 3, Appl
26	40.2	3.3	2628	US-08-696-944-1	Sequence 1, Appl
27	39.8	3.3	296	US-09-032-684-13	Sequence 13, Appl

28	39.8	3.3	1610	4	US-09-013-896A-1	Sequence 1, Appl
29	39.6	3.3	289	4	US-09-007-005-17	Sequence 17, Appl
30	39.6	3.3	289	4	US-09-244-796-17	Sequence 17, Appl
31	39.6	3.3	801	2	US-08-770-379-16	Sequence 16, Appl
32	39.6	3.3	801	4	US-08-757-669A-16	Sequence 16, Appl
33	39.6	3.3	801	4	US-09-230-371A-16	Sequence 16, Appl
34	39.6	3.3	1622	2	US-09-014-969-16	Sequence 16, Appl
35	39.2	3.2	470	4	US-09-020-956-102	Sequence 102, App
36	39.2	3.2	470	4	US-09-030-607-102	Sequence 102, App
37	39.2	3.2	470	4	US-09-439-313-102	Sequence 102, App
38	39.2	3.2	509	4	US-09-030-607-202	Sequence 202, App
39	39.2	3.2	509	4	US-09-439-313-202	Sequence 202, App
40	39.2	3.2	672	1	US-07-847-010-21	Sequence 21, Appl
41	39.2	3.2	1700	2	US-08-897-340-4	Sequence 4, Appl
42	39.2	3.2	1700	3	US-09-252-329-4	Sequence 4, Appl
43	39.2	3.2	1882	4	US-09-370-253-1	Sequence 1, Appl
44	39.2	3.2	1991	2	US-08-415-593-40	Sequence 40, Appl
45	39.2	3.2	2285	2	US-08-967-101-136	Sequence 136, App

ALIGNMENTS

RESULT 1
US-09-153-310-41
; Sequence 41, Application US/09153310
; Patent No. 6326184
; GENERAL INFORMATION:
APPLICANT: Gjernansen, Claes
Hansen, Jorgen
Johannesen, Pia Francke
Pedersen, Mogens Bohl
Sorensen, Steen Bech
TITLE OF INVENTION: Method of producing a composite
fermented beverage using genetically modified yeast
strains
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W.
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/153,310
FILING DATE: 15-Sep-1998
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: <Unknown>
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Bent, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: <Unknown>
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-672-5300
TELEFAX: 202-672-5399
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 41:
SEQUENCE CHARACTERISTICS:
LENGTH: 1160 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: genomic DNA
SEQUENCE DESCRIPTION: SEQ ID NO: 41:
US-09-153-310-41


```
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)856-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: ptzgpt-Fls
; US-08-232-463-14

Query Match 3.9%; Score 47.8; DB 1; Length 7218;
Best Local Similarity 4.0%; Pred. No. 0.0097;
Matches 16; Conservative 217; Mismatches 164; Indels 0; Gaps 0;

QY 301 TCCGGCTACGACTGCCACGCGGCATTTGGCGGTGGTGGCGCGCGCGGCGGAATGG 360
Db 1468 TCGAAGTAGTAAAGATAGAGAAGATTGGTACRNRNRNRNRNRNRNRNRNRNRNR 1409

QY 361 ACACAGCGCGGAGCCGCCACAGCCAGTGAAGGAGAACCTGTAAATCGAACAT 420
Db 1408 RNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNR 1349

QY 421 TCGGAAATCGACTAATTTATGGCACAATGCTTGTGGACAATCTGATAGACAAA 480
Db 1348 RNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNR 1289

QY 481 ATTGCTGGCAAAAGCGTGTGCTGTATGAGTAAACAGACTCAGTGTTCAGGAGAAA 540
Db 1288 RNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNR 1229

QY 541 TACTCTGCATGTGCATGAGTGTGAGTTCATTGACAGAGCCACCTCACGTATGTACT 600
Db 1228 RNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNR 1169

QY 601 TGATGGTGACAACTCAGACATGGCGCTAAATAGAGATTAAAGCTTTAAGCAGACG 660
Db 1168 RNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNR 1109

QY 661 TCGCAAAATATACGAGAGTGTGGTCAAGTGCAGAAAG 697
Db 1108 RNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNR 1072
```

```
RESULT 4
US-09-413-452-3
; Sequence 3, Application US/09413452
; Patent No. 6083540
; GENERAL INFORMATION:
; APPLICANT: Christensen, T.
; APPLICANT: Thorsoe, H.
; APPLICANT: Kreiberg, J.
; APPLICANT: Buchholt, H.
; TITLE OF INVENTION: PROCESS FOR STABILIZING PROTEINS IN AN
; FILE REFERENCE: DY0U14.001APC
```

```
; CURRENT APPLICATION NUMBER: US/09/413,452
; CURRENT FILING DATE: 1999-10-06
; EARLIER APPLICATION NUMBER: 08/983364
; EARLIER FILING DATE: 1998-05-18
; EARLIER APPLICATION NUMBER: PCT/EP96/03051
; EARLIER FILING DATE: 1996-07-12
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Citrus sinensis var. Navel
; US-09-413-452-3

Query Match 3.7%; Score 45.6; DB 3; Length 1323;
Best Local Similarity 52.7%; Pred. No. 0.017;
Matches 99; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 1030 TTAGTATATGTATTTGAGAGATTGATCTGATCTCTGTGTCCATTACTTGTGGACAC 1089
Db 1126 TTAATATGTTTGTATTAGTGGATCCATAGGTCTCTGGTCTTTCAATTGTGTAATTTGA 1185

QY 1090 AATAAGATCTGTTGTGTACATGAATAAAGGCGATCAACATGTAGCAAGAAC 1149
Db 1186 TTGAGCGTCTTATTCTGCTTCGATTTCACAAATACTATTGTGTGATTAAACAGAAA 1245

QY 1150 GTACGGTTCATTAGAAACGGATATGGATTTCATTCGTTTAAAAAATAAAAAA 1209
Db 1246 TAAATATGATCGGGAAGAATAAATTTCCGGCTTCTTTAAAAAATAAAAAA 1305

QY 1210 AAAAAAAA 1217
Db 1306 AAAAAAAA 1313

RESULT 5
US-09-413-068-3
; Sequence 3, Application US/09413068
; Patent No. 6268195
; GENERAL INFORMATION:
; APPLICANT: Christensen, T.
; APPLICANT: Thorsoe, H.
; APPLICANT: Kreiberg, J.
; APPLICANT: Buchholt, H.
; TITLE OF INVENTION: ACIDIC ENVIRONMENT WITH A HIGH-ESTER PECTIN
; FILE REFERENCE: DY0U14.001APC
; CURRENT APPLICATION NUMBER: US/09/413,068
; CURRENT FILING DATE: 1999-06-06
; EARLIER APPLICATION NUMBER: 08/983364
; EARLIER FILING DATE: 1998-05-18
; EARLIER APPLICATION NUMBER: PCT/EP96/03051
; EARLIER FILING DATE: 1996-07-12
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Citrus sinensis var. Navel
; US-09-413-068-3

Query Match 3.7%; Score 45.6; DB 4; Length 1323;
Best Local Similarity 52.7%; Pred. No. 0.017;
Matches 99; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 1030 TTAGTATATGTATTTGAGAGATTGATCTGATCTGTGTCCATTACTTGTGGACAC 1089
Db 1126 TTAATATGTTTGTATTAGTGGATCCATAGGTCTCTGGTCTTTCAATTGTGTAATTTGA 1185

QY 1090 AATAAGATCTGTTGTGTACATGAATAAAGGCGATCAACATGTAGCAAGAAC 1149
Db 1186 TTGAGCGTCTTATTCTGCTTCGATTTCACAAATACTATTGTGTGATTAAACAGAAA 1245
```



```
; FILING DATE: 05-MAY-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Leary Ph.D., Kathryn R.
; REGISTRATION NUMBER: 36,317
; REFERENCE/DOCKET NUMBER: DFCI-0029
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12001 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Herpes simplex virus
; STRAIN: Herpes Simplex Virus Type 1
; US-08-458-568A-11

Query Match 3.5%; Score 42.8; DB 1; Length 12001;
Best Local Similarity 51.6%; Pred. No. 0.24; Mismatches 0; Gaps 0;
Matches 98; Conservative 0; Indels 0;

QY 76 CCCAGCAGCGGCGGCAAGGCAACACCTCTCCAGCCGCGAGCGCTCGC 135
Db 1963 CCAAAACCTTCGGGGGCGCGCCGACCAACACCGCCCTCGCGCCCTCGCGCCCTCGC 2022
QY 136 CGTCATCTCTGAATCCACAGCGCGCTCCGCTCCAGGCTCACCCCTAGCGA 195
Db 2023 CCCCTCCGCGCCCTCGCCCTCCGCGCCCTCGCGCCCTCGCGCCCTCGCG 2082
QY 196 TCGGCACTCCGCGCGCTGTGATCATCGGCTACTCCCGGTTCCTACACTTCTTCGCG 255
Db 2083 CCCCTCGCGCCCTCCGCGCCCTCGCGCCCTCGCGCCCTCGCGCCCTCGCGCCCTCGCG 2142
QY 256 GGGTCTCGCC 265
Db 2143 CCTCCCGCC 2152

RESULT 9
US-09-165-264-11/c
; Sequence 11, Application US/09165264
; Patent No. 6197510
; GENERAL INFORMATION:
; APPLICANT: Vinayamoorthy, Thuralayah
; TITLE OF INVENTION: Multi-Loci Genomic Analysis
; FILE REFERENCE: 44747
; CURRENT APPLICATION NUMBER: US/09/165,264
; CURRENT FILING DATE: 1998-10-01
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 320
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Primer sequence
US-09-165-264-11

Query Match 3.5%; Score 42.4; DB 4; Length 320;
Best Local Similarity 49.1%; Pred. No. 0.06; Mismatches 116; Indels 0; Gaps 0;
Matches 112; Conservative 0;

QY 38 GTCACACACGACGACCAACACGAGCGCCGCGCCGCGCCGCGCCGCGCCGCGCCGCGCCG 97
Db 306 GGCACACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCC 247
QY 98 CAAGGCAACACCTCTCAGCGCCGACGCGCTCGCGTCTGTAATCCACAG 157
Db 246 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCC 187
QY 158 CGCGCGCGCTCCGCTCTCCAGGCTCACCCCTAGCGATCGCGCATCTCCGGCGCTCGTG 217
Db 186 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCC 127
QY 218 ATCATGGCTCACTCCCGGTTCCTCACTCTTCGCGGGTCTCGCC 265
Db 126 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCC 79

RESULT 11
US-08-474-379C-87/c
; Sequence 87, Application US/08474379C
; Patent No. 597305
; GENERAL INFORMATION:
; APPLICANT: Wigler, Michael H.
; APPLICANT: Colicelli, John J.
; TITLE OF INVENTION: CLONING BY COMPLEMENTATION AND RELATED
; TITLE OF INVENTION: PROCESSES
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive/6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```
Db 249 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCC 190
QY 150 ATCCACAGCGCGCGCTCTCCGCTCTCCAGGCTCACCCCTAGCGATCGCGCATCCCGG 209
Db 189 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCC 130
QY 210 CGCTCGTGATCATGCTCCCGGTTCCTCACTCTTCGCGGGTCTCGCC 265
Db 129 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCC 74

RESULT 10
US-09-165-264-14/c
; Sequence 14, Application US/09165264
; Patent No. 6197510
; GENERAL INFORMATION:
; APPLICANT: Vinayamoorthy, Thuralayah
; TITLE OF INVENTION: Multi-Loci Genomic Analysis
; FILE REFERENCE: 44747
; CURRENT APPLICATION NUMBER: US/09/165,264
; CURRENT FILING DATE: 1998-10-01
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 320
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Primer sequence
US-09-165-264-14

Query Match 3.5%; Score 42.4; DB 4; Length 320;
Best Local Similarity 49.1%; Pred. No. 0.06; Mismatches 116; Indels 0; Gaps 0;
Matches 112; Conservative 0;

QY 38 GTCACACACGACGACCAACACGAGCGCCGCGCCGCGCCGCGCCGCGCCGCGCCGCGCCG 97
Db 306 GGCACACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCC 247
QY 98 CAAGGCAACACCTCTCAGCGCCGACGCGCTCGCGTCTGTAATCCACAG 157
Db 246 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCC 187
QY 158 CGCGCGCGCTCCGCTCTCCAGGCTCACCCCTAGCGATCGCGCATCTCCGGCGCTCGTG 217
Db 186 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCC 127
QY 218 ATCATGGCTCACTCCCGGTTCCTCACTCTTCGCGGGTCTCGCC 265
Db 126 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCC 79

RESULT 11
US-08-474-379C-87/c
; Sequence 87, Application US/08474379C
; Patent No. 597305
; GENERAL INFORMATION:
; APPLICANT: Wigler, Michael H.
; APPLICANT: Colicelli, John J.
; TITLE OF INVENTION: CLONING BY COMPLEMENTATION AND RELATED
; TITLE OF INVENTION: PROCESSES
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive/6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,379C
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/511,715
FILING DATE: 20-APR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/206,188
FILING DATE: 01-MAR-1994
APPLICATION NUMBER: US 07/688,352
FILING DATE: 19-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 27866/32771
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
INFORMATION FOR SEQ ID NO: 87:
SEQUENCE CHARACTERISTICS:
LENGTH: 1481 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1008
US-08-474-379C-87

Query Match 3.5%; Score 42.4; DB 2; Length 1481;
Best Local Similarity 64.0%; Pred. No. 0.12;
Matches 64; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

Qy 1118 AAAAGGCATCAACATGTAGGAAGTACAGAGGTTCATTCAGAACGGATATGGA 1177
Db 1221 AAACACAGCAACCTGTATCCACTATTAGGAGGTAATAATCAATAAATGGCCCATTCAT 1162

Qy 1178 TTCATTCGTTTAAAAAATAAAAAAAAAAAAAAAAAAAAA 1217
Db 1161 TTGTGTGTAAAAAATAAAAAAAAAAAAAAAAAAAAA 1122

RESULT 12
US-07-688-352C-21/c
Sequence 21, Application US/07688352C
Patent No. 5527896
GENERAL INFORMATION:
APPLICANT: Wigler, Michael H.
APPLICANT: Colicelli, John J.
TITLE OF INVENTION: Cloning by Complementation and Related
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
ADDRESSEE: Bicknell
STREET: Two First National Plaza, 20 South Clark
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/688,352C

FILING DATE: 19910419
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/511,715
FILING DATE: 20-APR-1990
ATTORNEY/AGENT INFORMATION:
NAME: Borun, Michael F.
REGISTRATION NUMBER: 25447
REFERENCE/DOCKET NUMBER: 27805/30197
TELEPHONE: (312) 346-5750
TELEFAX: (312) 984-9740
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 3131 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1652
OTHER INFORMATION: /note= "A shift in reading frame
may occur at this residue."
FEATURE:
NAME/KEY: CDS
LOCATION: join(743..1648, 1651..2661)
US-07-688-352C-21

Query Match 3.5%; Score 42.4; DB 1; Length 3131;
Best Local Similarity 64.0%; Pred. No. 0.16;
Matches 64; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

Qy 1118 AAAAGGCATCAACATGTAGGAAGTACAGAGGTTCATTCAGAACGGATATGGA 1177
Db 2871 AAACACAGCAACCTGTATCCACTATTAGGAGGTAATAAATCAATAAATGGCCCATTCAT 2812

Qy 1178 TTCATTCGTTTAAAAAATAAAAAAAAAAAAAAAAAAAAA 1217
Db 2811 TTGTGTGTAAAAAATAAAAAAAAAAAAAAAAAAAAA 2772

RESULT 13
US-09-146-249A-21/c
Sequence 21, Application US/09146249A
Patent No. 6069240
GENERAL INFORMATION:
APPLICANT: Wigler, Michael H.
APPLICANT: Colicelli, John J.
TITLE OF INVENTION: Cloning by Complementation and Related
NUMBER OF SEQUENCES: 85
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/146,249A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/511,715
FILING DATE: 20-APR-1990
ATTORNEY/AGENT INFORMATION:


```

; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 743..1651
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1652
; OTHER INFORMATION: /note= "A shift in reading frame
; OTHER INFORMATION: may occur at this residue."
;
PCF-US91-02714-20

Query Match      3.5%; Score 42.4; DB 5; Length 3131;
Best Local Similarity 64.0%; Pred. No. 0.16;
Matches 64; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 1118 AAAAGGCATCAACATGTAGGAAGTAACACAGAGTACGGTTTCATTTCAGAAACGGATATGGA 1177
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2871 AAACCACAGCAACCTGTATCCACTATTAGGAGGTAAATAATCAATAAAATGCGCCCATTCAT 2812
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1178 TTCATTCGTTTAAAAAATAAAAAAAAAAAAAAAAAAAAAA 1217
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2811 TTGTGTTCTAAAAAATAAAAAAAAAAAAAAAAAAAAAA 2772
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: November 2, 2002, 01:42:23
Job time : 73 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2002, 23:00:32 ; Search time 1725 Seconds
(without alignments)
9522.200 Million cell updates/sec

Title: US-09-720-384A-3
Perfect score: 1217
Sequence: 1 ggcgcgtttcatttcac.....aaaaaaaaaaaaaaaaaaaaa 1217

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST.*
1: em_estha.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_htc.*
9: gb_estl.*
10: gb_est2.*
11: gb_htc.*
12: gb_gss.*
13: em_gss_hum.*
14: em_gss_inv.*
15: em_gss_pln.*
16: em_gss_vrt.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	590	48.5	675	10	BE357876 DGL_22_A0
2	583.4	47.9	691	10	BE355111 DGL_113_B
3	422.2	34.7	746	10	BE361884 DGL_82_H0
4	409.2	33.6	721	9	AW922946 DGL_47_E0
5	400.2	32.9	699	10	BE361874 DGL_82_G0
6	344.8	28.3	421	10	BF484142 WHE1788_G
7	330.6	27.2	627	9	AW052991 614077E08
8	318.8	26.2	591	10	BG873962 MEST44_F0
9	304.4	25.0	556	9	AW560397 EST315445
10	301.2	24.7	555	10	B1427055 sag10g11.
11	284.2	23.4	714	10	BM407099 EST581426
12	276.6	22.7	772	10	BF631366 HVSMEB001
13	275.2	22.6	748	10	BG126657 EST472303
14	271	22.3	576	10	BI320536 sah56h02.
15	268	22.0	564	9	AW056154 660004H03
16	265.6	21.8	272	9	AW288601 618069F05
17	261.6	21.5	786	10	B1178754 EST519699

18	259.6	21.3	596	9	AW508923
c 19	238	19.6	602	9	AI637166
c 20	233.4	19.2	503	9	AI987561
c 21	218.2	17.9	722	10	BE360898
c 22	216.4	17.8	560	10	BE404931
c 23	215.8	17.7	603	9	AW542370
c 24	208.4	17.1	631	9	AW279147
c 25	205.6	16.9	416	9	AW424749
c 26	200.6	16.5	530	10	BE823722
c 27	193.6	15.9	620	9	AW330778
c 28	171.6	14.1	809	9	BE195160
c 29	160.2	13.2	398	10	BF113246
c 30	159.6	13.1	839	9	AW983426
c 31	157.6	12.9	681	10	BI722646
c 32	156.2	12.8	745	10	BI935540
c 33	155.2	12.8	1069	12	CNS060UP
c 34	152.4	12.5	741	12	AQ365502
c 35	152.4	12.5	624	10	BI995324
c 36	152.2	12.5	387	10	BF113263
c 37	152.2	12.5	951	9	AL540583
c 38	152	12.5	886	9	AL550460
c 39	151.4	12.4	640	10	BI717399
c 40	151.4	12.4	712	10	BG852725
c 41	149.4	12.3	1041	10	BI829344
c 42	146	12.0	492	9	AW004219
c 43	139	11.4	737	10	BI768154
c 44	139	11.4	739	10	BI603869
c 45	139	11.4	746	9	AW957146

ALIGNMENTS

RESULT 1 BE357876 675 bp mRNA linear EST 20-JUL-2000
LOCUS DGL_22_A06.g1_A002 Dark Grown 1 (DGL) Sorghum bicolor cDNA, mRNA
DEFINITION BE357876 sequence.
ACCESSION BE357876
VERSION BE357876.1 GI:9299433
KEYWORDS EST.
SOURCE sorghum.
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE 1 (bases 1 to 675)
AUTHORS Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C., Sudman,M. and Pratt J.L.H.
TITLE An EST database from Sorghum: dark-grown seedlings
JOURNAL Unpublished (2000)
COMMENT Contact: Cordonnier-Pratt MM
The University of Georgia
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 542 1805
Email: mmpratt@uga.edu
Sequences have been trimmed to exclude PolyA, vector and regions below Phred quality 16. The threshold for highest quality sequence is 20.
Seq primer: PolyTMix
High quality sequence start: 30
High quality sequence stop: 638
POLYA-No. Location/Qualifiers
1. .675
/organism="Sorghum bicolor"
/db_xref="taxon:4558"
/clone_lib="Dark Grown 1 (DGL)"
/note="Organ: 5-day-old dark-grown seedlings; Vector: Lambda Zap; Site_1: XhoI; Site_2: EcoRI; The library was made from poly-A RNA in the cloning vector lambda Zap II.

BASE COUNT 207 a 121 c 161 g 186 t
ORIGIN

Query Match 48.5%; Score 590; DB 10; Length 675;
Best Local Similarity 93.6%; Pred. No. 1.6e-76;
Matches 640; Conservative 0; Mismatches 35; Indels 9; Gaps 2;

QY 466 ATCTGATAGACAGAAATTTGCTGGACAAAAGGCTGTGCTGATGATACAGGACTCAG 525
DB 1 ATCTGATAGACAGAAATTTGCTGGACAAAAGGCTGTGCTGATGATACAGGACTCAG 60
QY 526 TGGTTTCAGGAAAAGTACTCTTCATGTCGACTGAGTCGAGTTCGATGCGAGAGGCCA 585
DB 61 TGGTTTCAGGAAAAGTACTCTTCATGTCGACTGAGTCGAGTTCGATGCGAGAGGCCA 120
QY 586 CTTCACTGATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 645
DB 121 CATCACTGATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 180
QY 646 TAAGCAGAACCGCTCCACAAATATACGAAGAGTTGCTGAAGTGGCAAGCTTTTTCG 705
DB 181 TAAGCAGAACCGCTCCACAAATATACGAAGAGTTGCTGAAGTGGCAAGCTTTTTCG 240
QY 706 TGATGCTGCTCATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 765
DB 241 CGATGCTGCTCATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
QY 766 ATGCCGCTGCTTACTTCCACATCTTAACCTTTATGAGTATTTATGATTTGCCCTTAA 825
DB 301 ATGCCGCTGCTTACTTCCACATCTTAACCTTTATGAGTATTTATGATTTGCCCTTAA 360
QY 826 AATTTGTGAAGCTCGTGATCCTAAAGGCTATACAGGCTTGACAGTTCGACGAGAAATTA 885
DB 361 AATTTGTGAAGCTCGTGATCCTAAAGGCTTGACAGTTCGACGAGAAATTA 420
QY 886 AGGTTTCTACTGGAATTCATGATCCATACGAACCCCAATTAATGTTGAGATAGTAATTA 945
DB 421 AGGTTTCTACTGGAATTCATGATCCATACGAACCCCAATTAATGTTGAGATAGTAATTA 480
QY 946 GATGAAGATGAGAAATGCTTCACCCAAAGCAATGCCAAGAGTTCTATGCTACT 1005
DB 481 GATGAAGATGAGAAATGCTTCACCCAAAGCAATGCCAAGAGTTCTATCATACCT 540
QY 1006 TGAAGAAACCGATTTTGAAGCTTAGTATATATGTTTGAAGATTCGATCTGATCT 1065
DB 541 TGAAGAAACCGATTTTGAAGCTTAGTATATATGTTTGAAGATTCGATCTGATCT 597
QY 1066 TGTGTGTCATTTCTGTCGACACATAGATCTGTTGTTGTCACATGATTAAGGCA 1125
DB 598 TGTGTGTCATTTCTGTCGACACATAGATCTGTTGTTGTCACATGATTAAGGCA 651
QY 1126 TCAACATCTAGGAAGTAACAGAG 1149
DB 652 TCAACATCTAGGAAGTAACAGAG 675

RESULT 2
LOCUS BE355111
DEFINITION DGL113_B07.g1_A002 Dark Grown 1 (DGL1) Sorghum bicolor cDNA, mRNA
sequence.
ACCESSION BE355111
VERSION BE355111.1 GI:9296311
KEYWORDS EST.
SOURCE sorghum.
ORGANISM Sorghum bicolor

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Sorghum.
1 (bases 1 to 691)
Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C., Sudman,M. and Pratt
,L.H.

TITLE An EST database from Sorghum: dark-grown seedlings
JOURNAL Unpublished (2000)
COMMENT Contact: Cordonnier-Pratt MM
Department of Botany
The University of Georgia
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 542 1805
Email: mmpratt@uga.edu

Sequences have been trimmed to exclude polyA, vector and regions
below Phred quality 16. The threshold for highest quality sequence
is 20.

Seq primer: PolyTMix
High quality sequence start: 23
High quality sequence stop: 682
POLYA-No.

FEATURES
source

1. 691
/organism="Sorghum bicolor"
/db_xref="taxon:4558"
/clone_lib="Dark Grown 1 (DGL1)"
/note="Organ: 5-day-old dark-grown seedlings; Vector:
Lambda Zap; Site_1: XhoI; Site_2: EcoRI; The library was
made from poly-A RNA in the cloning vector lambda Zap II.
Clones to be sequenced were prepared by mass excision."
BASE COUNT 212 a 124 c 167 g 188 t
ORIGIN

Query Match 47.9%; Score 583.4; DB 10; Length 691;
Best Local Similarity 92.8%; Pred. No. 1.4e-75;
Matches 636; Conservative 0; Mismatches 41; Indels 8; Gaps 2;

QY 473 AGACAGAAATTCCTGGGACAAAAGGCTGTGCTATGATGATAACAGGACTCAGTGTTC 532
DB 1 AGACAGAAATTCCTGGGACAAAAGGCTGTGCTGATGATAACAGGACTCAGTGTTC 60
QY 533 GGCAGAAAGTACTTTCGATGTCACCTGAGTCGAGTTCGATGATGATGATGATGATGAT 592
DB 61 GGCAGAAAGTACTTTCGATGTCACCTGAGTCGAGTTCGATGATGATGATGATGATGAT 120
QY 593 TATGTACTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 652
DB 121 TATGTACTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
QY 653 GAAGACCGTGCAGAAAATATACGAAGAGTTGTTGAAGTGGCAAGCTTTTTCGATGCT 712
DB 181 GAAGACCGTGCAGAAAATATACGAAGAGTTGTTGAAGTGGCAAGCTTTTTCGATGCT 240
QY 713 GGTGTCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 772
DB 241 GGTATCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
QY 773 GCTCTACTTCCACATTCCTAACTTTTATGAAGTATTTTATGATTTGCCCTTAAATTTGT 832
DB 301 GCTCTACTTCCACATTCCTAACTTTTATGAAGTATTTTATGATTTGCCCTTAAATTTGT 360
QY 833 GAAGCTCGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 892
DB 361 GAAGCTCGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
QY 893 ACTGGAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 952
DB 421 ACTGGAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
QY 953 GATGAGGAATGCTTTCACCCAAAGCAATGCGCAAGCTTCATGCTACCTTTGAAGAA 1012
DB 481 GATGAGGAATGCTTTCACCCAAAGCAATGCGCAAGCTTCATGCTACCTTTGAAGAA 540
QY 1013 AAGCGATTTTGAAGCTTTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1072
DB 541 AAGCGATTTTGAAGCTTTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 597
QY 1073 CCATTACTTGTGGACACAAATAAGATCTGTTGTTGGTCACATGATTAAGAGCATCAACAT 1132


```

Db 598 CCATTACTTGTGACACAATGATCTGTTG-----GCCATGAATAAAGGCATCAACAC 652
      |||
QY 1133 GTAGGAAGTAACAGAAGTACGGTT 1157
      |||
Db 653 ATAGCAAGTAACAGAGCGCGAGTTT 677
      |||

RESULT 3
BE361884
LOCUS      BE361884
DEFINITION BE361884 746 bp mRNA linear EST 20-JUL-2000
            Dg1_82_H05_g1_A002 Dark Grown 1 (Dg1) Sorghum bicolor cDNA, mRNA
            sequence.
ACCESSION  BE361884
VERSION     BE361884
KEYWORDS    BE361884.1 GI:9303441
SOURCE      EST.
ORGANISM    sorghum.
            Sorghum bicolor
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE   1 (bases 1 to 746)
AUTHORS     Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C., Sudman,M. and Pratt
            ,L.H.
            An EST database from Sorghum: dark-grown seedlings
            Unpublished (2000)
            Contact: Cordonnier-Pratt MM
            Department of Botany
            The University of Georgia
            Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
            Tel: 706 542 1860
            Fax: 706 542 1805
            Email: mmpratt@uga.edu
            Sequences have been trimmed to exclude PolyA, vector and regions
            below Phred quality 16. The threshold for highest quality sequence
            is 20.
            Seq primer: PolyTMix
            High quality sequence start: 30
            High quality sequence stop: 719
            POLYA-No. Location/Qualifiers
            source      1..746
                       /organism="Sorghum bicolor"
                       /db_xref="taxon:4558"
                       /clone_lib="Dark Grown 1 (Dg1)"
                       /note="Organ: 5-day-old dark-grown seedlings; Vector:
                       Lambda Zap; Site.1: XhoI; Site.2: EcoRI; The library was
                       made from poly-A RNA in the cloning vector lambda ZAP II.
                       Clones to be sequenced were prepared by mass excision."
BASE COUNT  218 a 126 c 166 g 236 t
ORIGIN
Query Match 34.7%; Score 422.2; DB 10; Length 746;
Best Local Similarity 90.3%; Pred. No. 3.2e-52;
Matches 476; Conservative 0; Mismatches 43; Indels 8; Gaps 2;

QY 664 ACAAATATACGAAGCTTGGTGAAGTGGCAAGCTTTTTCCTCATGCTGGTGCATATG 723
      |||
Db 1 AGAAATATACGAAGCTTGGTGAAGTGGCAAGCTTTTTCCTCATGCTGGTGCATATG 60
      |||
QY 724 CATTGCTAGCTTGATATCTCCATACAGGAGAGATCGTGATGCGGTGCTCTACTTCC 783
      |||
Db 61 CATTGCTAGCTTGATATCTCCATACAGGAGAGATCGTGATGCGGTGCTCTACTTCC 120
      |||
QY 784 ACATTCTTAACCTTATTGAAGTATTATTGATTGGCCCTAAAATTTGTGAAGCTCGTGA 843
      |||
Db 121 AGATTCTAACCTTATTGAAGTATTATTGATTGGCCCTAAAATTTGTGAAGCTCGTGA 180
      |||
QY 844 TCCTAAAGCCCTATACAAAGCTTCGACGTACAGGAAGATTAAAGGTTTCACTGGGAATTGA 903
      |||
Db 181 TCCTAAAGGCTGTGTACAAGCTTCGACGCACAGGAAGATTAAAGGTTTCACTGGGAATTGA 240
      |||
QY 904 TGATCCATACGAACCAATTAAATGGTGAGATAGTAATTAAGATGAAGATGAGGAATG 963
      |||

```

```

Db 241 TGATCCATACGAACCGCCAGTTAATGGTGAGATAGTAATTAGATGAAGATGGGAATG 300
      |||
QY 964 CCTTTCACCCCAAGCAATAGCCCAAGTTCCTATGCTACCTTGAAGAAACGGATATTT 1023
      |||
Db 301 CGCTTCACCCCAAGCAATAGCCCAAGTTCCTATGCTACCTTGAAGAAACGGATATTT 360
      |||
QY 1024 GCAAGCTTAGTATATATTGAGAGAGATGATCTGATGATCTTGTGTGTCATTACTTGT 1083
      |||
Db 361 GCAAGCTTAGCACAT--ATTGAGATATGATCTGATCTTGTGTGTCATTACTTGT 417
      |||
QY 1084 GGACACAATAAGATCTGTGTGTCACATGAATAAAGGCATCAACATGTAGGAAGTAA 1143
      |||
Db 418 GGACACAATAATGATCTGTTG-----GCCATGAATAAAGGCATCAACACATAGCAAGTAA 472
      |||
QY 1144 CAGAAGGTACGGTTCATTCAGAAACGGATGATGATTTCATTGTTTAA 1190
      |||
Db 473 CAGAGCGCAGTTGTTTCAGAAACGGGAATACAGATTCATTGCTTCA 519
      |||

RESULT 4
AW922946
LOCUS      AW922946
DEFINITION Dg1_47_E04_g1_A002 Dark Grown 1 (Dg1) Sorghum bicolor cDNA, mRNA
            sequence.
ACCESSION  AW922946
VERSION     AW922946
KEYWORDS    AW922946.1 GI:8088771
SOURCE      EST.
ORGANISM    sorghum.
            Sorghum bicolor
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE   1 (bases 1 to 721)
AUTHORS     Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C., Sudman,M. and Pratt
            ,L.H.
            An EST database from Sorghum: dark-grown seedlings
            Unpublished (2000)
            Contact: Cordonnier-Pratt MM
            Department of Botany
            The University of Georgia
            Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
            Tel: 706 542 1860
            Fax: 706 542 1805
            Email: mmpratt@uga.edu
            Sequences have been trimmed to exclude PolyA, vector and regions
            below Phred quality 16. The threshold for highest quality sequence
            is 20.
            Seq primer: PolyTMix
            High quality sequence start: 23
            High quality sequence stop: 703
            POLYA-No. Location/Qualifiers
            source      1..721
                       /organism="Sorghum bicolor"
                       /db_xref="taxon:4558"
                       /clone_lib="Dark Grown 1 (Dg1)"
                       /note="Organ: 5-day-old dark-grown seedlings; Vector:
                       Lambda Zap; Site.1: XhoI; Site.2: EcoRI; The library was
                       made from poly-A RNA in the cloning vector lambda ZAP II.
                       Clones to be sequenced were prepared by mass excision."
BASE COUNT  202 a 124 c 161 g 234 t
ORIGIN
Query Match 33.6%; Score 409.2; DB 9; Length 721;
Best Local Similarity 90.1%; Pred. No. 2.5e-50;
Matches 463; Conservative 0; Mismatches 43; Indels 8; Gaps 2;

QY 677 AGAGTTGGTGAAGTGGCAAGCTTTTTCCTGATGCTGGTGCATATGCTAGCTTGT 736
      |||
Db 1 AGAGTTGGTGAAGTGGCAAGCTTTTTCCTGATGCTGGTGCATATGCTAGCTTGT 60
      |||
QY 737 ATATCTCCATACAGGAGAGATCGTGATGCGGTGCTCTACTTCCACATCTTAACCTT 796
      |||

```

Db 61 ATATCTCCATACAGAGAGATCGTGATGATGCGCGTCTACTTCCAGATTCTAACTTT 120
QY 797 ATTGAAGTATTTATTGATTTGCCCTTAAATAATTTGTGAAGCTCGTGATCCTAAAGGCCCTA 856
Db 121 ATTGAAGTATTTATTGATTTGCCCTTAAATAATTTGTGAAGCTCGTGATCCTAAAGGCTG 180
QY 857 TACAAGCTTGACACCTACAGAAAGATTAAAGTTTCACTGGAATTCATGATCCATACGAA 916
Db 181 TACAAGCTTGACACCTACAGAAAGATTAAAGTTTCACTGGAATTCATGATCCATACGAA 240
QY 917 CCACCAATTAATGTGAGATGATGATTAAGATGAAGATGAGGAATGCCCTTACACCAA 976
Db 241 CGCGCTGATTAATGTGAGATGATGATTAAGATGAAGATGAGGAATGCCCTTACACCAA 300
QY 977 GCAATGGCCAAAGCAAGTTCTATGCTACCTTGAAGAAACGGATATTGCAAGCTTAGPAT 1036
Db 301 GCAATGGCCAAAGCAAGTTCTATGCTACCTTGAAGAAACGGATATTGCAAGCTTAGPAT 360
QY 1037 ATGATTTTGTGAGAGATGATGATGATTTCTGTTGTCATCTTGTGACACAAATAGA 1096
Db 361 AT---ATTGTGAGATGATGATGATTTCTGTTGTCATCTTGTGACACAAATAGA 417
QY 1097 TCTGTTGTTGTCATGATTAAGGATCAACATGATGAGGAATTAACAGAGTACGGT 1156
Db 418 TCTGTTG-----GCCATGAATAAAGGATCAACATAGCAAGTAAACAGAGGCGCAGTT 472
QY 1157 TCATTCAGAAACGGATGATTCATTCGTTAA 1190
Db 473 TGTTCAGAAACGGATACAGATTCATTCGTTCA 506

RESULT 5
BE361874
LOCUS
DEFINITION
DGI_82.G05.g1_A002 Dark Grown 1 (DGI) Sorghum bicolor cDNA, mRNA
EST 20-JUL-2000
sequence.
ACCESSION
BE361874.
VERSION
BE361874.1 GI:9303431
KEYWORDS
EST.
SOURCE
sorghum.
ORGANISM
Sorghum bicolor

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Sorghum.
1 (bases 1 to 699)
Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C., Sudman,M. and Pratt
L.H.
TITLE
An EST database from Sorghum: dark-grown seedlings
JOURNAL
Unpublished (2000)
COMMENT
Contact: Cordonnier-Pratt MM
Department of Botany
The University of Georgia
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 542 1805
Email: mmpratt@uga.edu

Sequences have been trimmed to exclude PolyA, vector and regions
below Phred quality 16. The threshold for highest quality sequence
is 20.
Seq primer: PolyTmix
High quality sequence start: 64
High quality sequence stop: 697
POLYA=No.
Location/Qualifiers

source
1. .699.
/organism="Sorghum bicolor"
/db_xref="taxon:4558"
/clone_lib="Dark Grown 1 (DGI)"
/note="Organ: 5-day-old dark-grown seedlings; Vector:
Lambda Zap; Site_1: XhoI; Site_2: EcoRI; The library was
made from poly-A RNA in the cloning vector Lambda ZAP II.
Clones to be sequenced were prepared by mass excision."

BASE COUNT 199 a 121 c 154 g 225 t
ORIGIN
Query Match 32.9%; Score 400.2; DB 10; Length 699;
Best Local Similarity 89.9%; Pred. No. 5.1e-49;
Matches 454; Conservative 0; Mismatches 43; Indels 8; Gaps 2;
QY 686 GAAGTGGCAAGCAAGCTTTTGGTGTGATGCTGGTGTGCATATGCATTGCTAGCTTGAATCTCCA 745
Db 1 GAAGTGGCAAGCAAGCTTTTGGCGGATGCTGGTATCATATGCATTGCTAGCTTGAATCTCCA 60
QY 746 TACAGGAGATCGTGATGATGCCGTGCTCTACTTCCACATTTCTAACTTTATTGAAGTA 805
Db 61 TACAGGAGATCGTGATGATGCCGTGCTCTACTTCCAGATTTCTAACTTTATTGAAGTA 120
QY 806 TTTATTGATTTGCCCTTAAATAATTTGTGAAGCTCGTGATCCTAAAGGCTTATACAAGCTT 865
Db 121 TTTATTGATTTGCCCTTAAATAATTTGTGAAGCTCGTGATCCTAAAGGCTTATACAAGCTT 180
QY 866 GCACGTACAGGAAGATTAAGGTTTCACTGGAATTTGATGATCCATACGAAACCAATTT 925
Db 181 GCACGTACAGGAAGATTAAGGTTTCACTGGAATTTGATGATCCATACGAAACCAATTT 240
QY 926 AATGTTGATGATTAAGATCAAGATGAGGAATGCCCTTCAACCAAGCAATGGCC 985
Db 241 AATGTTGATGATTAAGATCAAGATGAGGAATGCCCTTCAACCAAGCAATGGCC 300
QY 986 AAGCAAGTTCTGCTACCTTGAAGAAACGGATTTTGAAGCTTAGTATGATGATTTT 1045
Db 301 AAGCAAGTTCTGCTACCTTGAAGAAACGGATTTTGAAGCTTAGTATGATGATTTT 357
QY 1046 GAGAAGATGATGATTTGTTGTTCCATTACTTTGGGACACAATAGATGTTGTT 1105
Db 358 GAGAATATGATGATTTGTTGTTCCATTACTTTGGGACACAATAGATGTTGTTG-- 415
QY 1106 GGTTCATGATTAAGGCAATCAACATGATGAGGAATGAGGAATGAGGATGAGGTTTCATTGAGA 1165
Db 416 ---GCCATGATTAAGGCAATCAACATGATGAGGAATGAGGATGAGGATGAGGTTTCATTGAGA 472
QY 1166 AACGATATGATGATTTGTTGTTTAA 1190
Db 473 ACGGAATACAGATTCATTCGTTTCA 497

RESULT 6
BF484142
LOCUS
DEFINITION
WHE1788_G10_M202S wheat pre-anthesis spike cDNA library Triticum
aestivum cDNA clone WHE1788_G10_M20, mRNA sequence.
ACCESSION
BF484142
VERSION
BF484142.1 GI:11567443
KEYWORDS
EST.
SOURCE
bread wheat.
ORGANISM
Triticum aestivum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae
; Triticeae; Triticum.
1 (bases 1 to 421)
Anderson,O.D., Chao,S., Choi,D.W., Close,T.J., Fenton,R.D., Han
P.S., Hsia,C.C., Kang,Y., Lazo,G.R., Miller,R., Rausch,C.J.,
Seaton,C.L. and Tong,J.C.
The structure and function of the expressed portion of the wheat
genomes - Pre-anthesis spike cDNA library
Unpublished (2000)
Contact: Olin Anderson

US Department of Agriculture, Agriculture Research Service, Pacific
West Area, Western Regional Research
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105595773
Fax: 5105595818
Email: oanderson@w.usda.gov
Sequence have been trimmed to remove vector sequence and low
quality sequence with phred score less than 20

	Clade: Panicoideae; Andropogoneae; Zea.					
REFERENCE	1 (bases 1 to 627)					
AUTHORS	Walbot,V.					
TITLE	Maize ESTs from various cDNA libraries sequenced at Stanford University					
JOURNAL	Unpublished (1999)					
COMMENT	Contact: Walbot V Department of Biological Sciences Stanford University 855 California Ave, Palo Alto, CA 94304, USA Tel: 650 723 2227 Fax: 650 725 8221 Email: walbot@stanford.edu Plate: 614077 row: E column: 08.					
FEATURES						
SOURCE	Location/Qualifiers 1..627 /organism="Zea mays" /cultivar="W23" /db_xref="taxon:4577" /clone_lib="614 - root cDNA library from Walbot Lab" /tissue_type="root" /dev_stage="3-4 days old" /lab_host="XL0LR" /note="Organ: root; Vector: pBlueScriptII SK+; Site:1: ECORI; Site:2: XhoI; 3-4 days old root tissue from Walbot Lab (LM)"; 190 a 143 c 111 g 182 t 1 others					
BASE COUNT	Query Match 27.2%; Score 330.6; DB 9; Length 627; Best Local Similarity 96.9%; Pred. No. 6.6e-39; Matches 348; Conservative 0; Mismatches 9; Indels 2; Gaps 1					
ORIGIN						
Qy	819	CCCTAAATTTGTGAGCTCGTGATCCTTAAGGCCTATAACAGCTTGCACGTACAGGAA	878			
Dd	625	CCCCAAAATTGTGAGCTCGTGATCCTTAAGGCCTATAACAGCTTGCACGTACAGGAA	566			
Qy	879	AGATTTAAGGGTTTCACCTGGAATTTGATGATCCATAGCAACCACCAATTAATGTGAGATG	938			
Dd	565	AGATTTAAGGGTTTCACCTGGAATTTGATGATCCATAGCAACCACCAATTAATGTGAGATG	506			
Qy	939	TAATTAAGATGAAGATGAGGAATGCCCTTACCCAAAGAACAATGGCCAAAGTTCAT	998			
Dd	505	TAATTAAGATGAAGATGAGGAATGCCCTTACCCAAAGAACAATGGCCAAAGTTCAT	446			
Qy	999	GCTACTTGAAGAAACGGATATTTGCAAGCTTAGTATATGTTTGAAGAGATTGATC	1058			
Dd	445	GCTACTTGAAGAAACGGATATTTGCAAGCTTAGTATATGTTTGAAGAGATTGATC	388			
Qy	1059	TGATTCCTGTGTGCCATTACTTGTGGACACAATAAGATCTGTGTGGTCCACATGAATA	1118			
Dd	387	TGATTCCTGTGTGCCATTACTTGTGGACACAATAAGATCTGTGTGGTCCACATGAATA	328			
Qy	1119	AAGGCATCACATGTAGGAAGTATACAGAAGTAGCGGTTCAATTCAGAACGGATATGA	1177			
Dd	327	AAGGCATCACATGTAGGAAGTATACAGAAGTAGCGGTTCAATTCAGAACGGATATGA	269			
RESULT	8					
LOCUS	BG873962/C					
DEFINITION	MES14-F08.T3 ISUM4-TN Zea mays cDNA clone MES14-F08 3', mRNA sequence.					
ACCESSION	BG873962					
VERSION	BG873962.1 GI:14245380					
KEYWORDS	EST.					
SOURCE	Zea mays.					
ORGANISM	Zea mays.					
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.					
AUTHORS	Qiu,F., Cui,F., Guo,L., Ashlock,D.A, Wen,T.J. and Schnable,P.S.					

TITLE Expressed Sequence Tags from B73 Maize Seedlings and Silks
JOURNAL Unpublished (2001)
COMMENT Contact: Patrick S. Schnable
 Schnable Laboratory
 Iowa State University
 G405 Agronomy, Iowa State University, Ames, IA 50011-1010, USA
 Tel: 515-294-0975
 Fax: 515-294-2299
 Email: schnable@iastate.edu

PCR Primers
FORWARD: T7-1 (AA TAC GAC TCA CTA TAG)
BACKWARD: T3 (ATT AAC CCT CAC TAA AG)
Seq primer: primer T3 (ATT AAC CCT CAC TAA AG).

FEATURES

1..591
 Location/Qualifiers
 .organism="Zea mays"
 .cultivar="B73"
 .db_xref="taxon:4577"
 .clone="MESF44-F08"
 .clone_lib="ISUM4-TN"
 .tissue_type="Seedling and silk"
 .lab_host="DH10B"
 /note="Vector: pT73PAC; Site_1: EcoRI; Site_2: NotI;
 ds-cDNA molecules were generated as follows. First-strand
 cDNA was prepared from oligo-dT selected mRNA by priming
 with a NotI oligo-dT primer (5'
 AACTGGAACATTCGCGCGCAGGAATTTTTTTTTTTT). The
 resulting DNA-RNA hybrid was treated with RNase H and used
 as a template for DNA Pol-I-catalyzed second strand
 synthesis. After the addition of EcoRI adaptors, the
 ds-cDNAs were digested with NotI and size-selected. The
 resulting molecules were directionally cloned into the
 EcoRI and NotI sites of the pT73PAC vector. The library
 then went through one round of normalization to Cot value
 of 5 based on the methods of Marcelo Bento Soares (Genome
 Research 6: 791-806, 1996)."
 Research 6: 791-806, 1996)."

BASE COUNT 183 a 130 c 95 g 183 t
ORIGIN

Query Match 26.2%; Score 318.8; DB 10; Length 591;
 Best Local Similarity 97.9%; Pred. No. 3.5e-37;
 Matches 323; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 861 AGCTTGACGTACAGGAAGATTAAAGTTTCACTGGAATTGATGATCCATACGAACCA 920
 Db 591 AGCTTGACGTACAGGAAGATTAAAGTTTCACTGGAATTGATGATCCATACGAACCA 532
 Qy 921 CAATTAATGTTGACATAGTAATTAAGATGAAGATGAGGATGCCCTTCAACCAAGCAA 980
 Db 531 CAATTAATGTTGACATAGTAATTAAGATGAAGATGAGGATGCCCTTCAACCAAGCAA 472
 Qy 981 TGCCCAAGCAAGTTCTATGCTACCTTGAAGAAACGGATATTTGCCAAGCTTAGTATGT 1040
 Db 471 TGCCCAAGCAAGTTCTATGCTACCTTGAAGAAACGGATATTTGCCAAGCTTAGTATGT 412
 Qy 1041 ATTTTGACAAGATTGATCTGTGTGTCCTACTTGTGACACAAATGAAGTCG 1100
 Db 411 ATTTTGACAAGATTGATCTGTGTGTCCTACTTGTGACACAAATGAAGTCG 352
 Qy 1101 TTGTTGGTCACATGAATAAAGGCATCAACATGTAGGAAGTAACAGAAGTCAGGTTTAT 1160
 Db 351 TTGTTGGTCACATGAATAAAGGCATCAACATGTAGGAAGTAACAGAAGTCAGGTTTAT 292
 Qy 1161 TCAGAAACGGATGGATTTCGTTTAA 1190
 Db 291 TCAGAAACGGATGGATTTCGTTTAA 262

RESULT 9
AW560397
LOCUS AW560397 536 bp mRNA linear EST 07-SEP-2000
DEFINITION EST315445 DSIR Medicago truncatula cDNA clone pDSIR-27A7, mRNA sequence.

ACCESSION

VERSION
 KEYWORDS
 SOURCE
 ORGANISM

AW560397
 AW560397.1 GI:7205823
 EST.
 barrel medic.
 Medicago truncatula

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
 Medicago.

1 (bases 1 to 536)

Fedorova,M., Pierson,B.L., Samac,D.A., Vance,C.P., Gantt,G.S., Peng
 H., Ellis,L., Town,C.D., Bowman,C.L., Craven,M.B., Hansen,T.S.,
 Holt,I.E. and Fraser,C.M.

ESTs from roots of Medicago truncatula after inoculation with

Phytophthora medicaginis

Unpublished (1999)

Contact: Carroll P. Vance

Department of Agronomy and Plant Genetics

University of Minnesota

411 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108 USA

Tel: 612 625 5715

Fax: 651-649-5058

Email: vance004@maroon.tc.umn.edu

Minnesota EST name:M251752e ; TIGR sequence name:MTBAY04TK ; More

information, including clone ordering, is available at.

'http://chrysis.tamu.edu/medicago'

Seq primer: SKmod (CTA gAA CTA gtg gAT CC).

Location/Qualifiers

1..536

/organism="Medicago truncatula"

/cultivar="genotype A17"

/db_xref="taxon:3880"

/clone="pDSIR-27A7"

/clone_lib="DSIR"

/tissue_type="roots infected with Phytophthora

medicaginis"

/dev_stage="roots harvested at 10 days post inoculation

with Phytophthora medicaginis"

/lab_host="E. coli strain XL0LR"

/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
 XhoI; cDNA was prepared from polyA+ enriched RNA from
 roots harvested at 10 days post inoculation with
 Phytophthora medicaginis. The cDNA was directionally
 ligated into the Uni-ZAP XR vector from Stratagene and
 packaged using Gigapack III Gold packaging extracts.
 Plasmids containing cDNA inserts were excised from the
 recombinant lambda-ZAP phage using Ex-Assist helper phage
 and propagated in XL0LR cells. Note: EST may be of fungal
 origin."

BASE COUNT 156 a 102 c 130 g 148 t

ORIGIN

Query Match 25.0%; Score 304.4; DB 9; Length 536;
 Best Local Similarity 74.3%; Pred. No. 4.4e-35;
 Matches 397; Conservative 0; Mismatches 136; Indels 1; Gaps 1;

Qy 404 CCTGTAATGTCGACATTTGGAAATCGACTAATATTTATGGCACATTCCTTGATGGA 463
 Db 4 CCTGAGATGTCGAAATTTGGAACTCGACGAACATTTTGGGCATGATGTCCAGTTCAA 63
 Qy 464 CAATCTGATAGACAGAAATTTGCTGGGACAAAAGCGTGTGTCGTATGATGAACAGACTC 523
 Db 64 AAATGTGATACAGCAGCTGCTTCAGCAAAAAGGATGTGTATATGATGGCTAACTGCTC 123
 Qy 524 AGTGGTTCAGGAAAAGTACTCTTCATCTGCATGCTGCTGATGTTGATGTCGAGAGGC 583
 Db 124 AGTGGTTCAGGAAAAGTACTCTTCATCTGCATGCTGCTGATGTTGATGTCGAGAGGC 183
 Qy 584 CACCTCAGTATGATCTTGTGATGTCACACCTTCAGACATGGCCCTTAATAGATTTAAGC 643
 Db 184 AAATGACTTACATCTCTTGCGGTGACATAATTCGGCATGCTCTAAACCGCTGATCTTAGT 243
 Qy 644 TTTAAGCAGGAAGACCGCTGCGAGAAATATACGAGAGTTGGTGAAGTGGCGAAAGCTTTT 703

BM407099
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source
Location/Qualifiers
1. .714
/organism="Solanum tuberosum"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="cPRQ29H6"
/clone_lib="potato roots"
/tissue_type="roots"
/dev_stage="in vitro grown stem cuttings"
/lab_host="SOLR"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; supplier: Cornell University, Tanksley lab; sequencing: The Institute for Genomic Research. Roots were isolated from in vitro grown stem cuttings on CM medium. Roots were isolated two weeks after placing the stem cuttings from in vitro grown plants on medium."
BASE COUNT
ORIGIN
Query Match
Best Local Similarity 23.4%; Score 284.2; DB 10; Length 714;
Mismatches 0; Conservative 0; Mismatches 158; Indels 0; Gaps 0;
QY 408 TAATCTGCAACATGGCAATCGACTATATTTATGGCACATTCCTTGATGGACAAT 467
Db
QY 468 CTGATPAGACAAATTCCTGGGACAAAAGGCTGTGCTGATGGATAACAGGACCTCAGTG 527
Db
QY 528 GTTCAGGGAAGTACTCTGTCATGTCGACTGCTGAGTTCGATTCGACAGGCCACC 587
Db
QY 588 TCACATGTACTTGTGTGATGTCGACACCTCAGACATGGCCCTAAATAGAGATTTTAAGCTTTA 647
Db
QY 648 AGGCAGGAAGCCGTGGCAGAAAATATACGAAGTTCGTGAAGTGCGCAAAAGCTTTTTCGTG 707
Db
QY 708 ATGCTGGTGTCATATGATTCGATTCGATTCATCCATACAGAGAGATCGTGATGCAT 767
Db
QY 768 GCGGTGCTCTACTTCACATCTTAACCTTTATTTGAAGTATTTATTTGATTTGGCCCTCAAAA 827

Db 417 GCCGCCCATTTATGCCAGATAAAAAGTTCATTCAGGTTTTTATGAATATGCCTCTACAAC 476
QY 828 TTTGTGAAGCTGTGATCCTCTAAAGGCTTATACAAGCTTGACAGTACAGGAAGATTAAAG 887
Db 477 TGTGTGAAGATAGATCCAAAGGCTCTACAAGCTAGCTGTGAGGTAAGATCAAG 536
QY 888 GTTTCACCTGGAATTGATGATCCATACGAACCAACCAATTAATGTTGATAGATTAATTA 944
Db 537 GTTTTACTGGAGTAGATGATCCTTATGAACCACTTTTGAATTTGTGAGGATGTATTA 593
RESULT 12
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source
1. .772
/organism="Hordeum vulgare"
/cultivar="Morex"
/clone="HVSMB0015L17f"
/clone_lib="Hordeum vulgare seedling shoot EST library
HVCDNA0002 (Dehydration stress)"
/tissue_type="Seedling shoot"
/lab_host="TJC121"
/note="Vector: lambdaZAP; Site_1: EcoRI; Site_2: XhoI;
Seeds were surface sterilized then germinated under axenic
conditions in the dark at room temperature on filter paper
with water, nystatin and cefotaxime in covered
crystallization dishes. Five-day old seedlings were
incubated at 90% RH for 24 hr. Shoots were then harvested,
total RNA was prepared, poly(A) RNA was purified, one
primary unamplified cDNA library was made, 60000 pfu were
in vivo excised to give pBluescript SK(-) cDNA phagemids.
These steps were performed in the TJ Close laboratory at
the University of California, Riverside (Choi, Close,
Fenton). Phagemids were plated and picked at the Clemson
University Genomics Institute (CUGI) (Begum, Palmer,
Frisch, Atkins and Wing). Plasmid DNA preparations, DNA
sequencing and sequence analysis were performed at CUGI
(Wing, Yu, Frisch, Henry, Simmons, Oates, Rambo, Main).
The sequence has been trimmed to remove vector sequence
and contains a minimum of 100 bases of phred value 20 or
above. For more details on library preparation and

sequence analysis see
http://www.genome.clemson.edu/projects/barley. To order
this clone see http://www.genome.clemson.edu/orders/Also
see Close TJ, Wing R, Kleinhofs A, Wise R (2001)
Genetically and physically anchored EST resources for
barley genomics. Barley Genetics Newsletter 31:29-30.
(http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html)*

BASE COUNT 202 a 146 c 238 g 186 t

Query Match 22.7%; Score 276.6; DB 10; Length 772;
Best Local Similarity 71.5%; Pred. No. 4.1e-31;
Matches 377; Conservative 0; Mismatches 149; Indels 1; Gaps 1;

QY 412 CTCGACAAATTCGGGAATCGACTAATATTTATGGCACAATTCGTTGATGGACATCTGA 471
DB 71 GTCACTGTGCCAAGTCATCAATATCTTCGGCATGATGCGGAGTGGCAAGACTGA 130
QY 472 TAGACAGAAATTCGTGGGACAAAGGCTGTGCTATGGATACAGGACTCAGTGGTTC 531
DB 131 CCGGCAAAATCTACTGAAGCAGAAAGGCTGTGTTTGGATTACAGGCTTAGTGGTTC 190
QY 532 AGGGAAGTACTCTTCGATGTCGACTGAGTCGTGAGTTGCATTCGAGAGGCCACTCAC 591
DB 191 AGTAAAAGTACCTTGGCATGACATAGGTCGAGAGCTCCATACAGAGGAAGCTTGC 250
QY 592 GTATGACTTGGTGGACAACTCAGACATGGGCTAAATAGAGATTTAAGCTTTAAGGC 651
DB 251 GTATGCTTGTGGGATACCTTAAGACATGGCTTTAACAAAGATCTTGGGTTGCGAGC 310
QY 652 AGAAGCGTCGCAAAATATACGAAGAGTTGGTGAAGTGGCAAGCTTTTTCGTGATGC 711
DB 311 TGAAGATCGTCTGAAAATATACCGAGAGTTGGTGAAGTGGCAAGCTATTGGCAGTGC 370
QY 712 TGGTGTATATGCTAGCTTGATATCTCCATACAGGAGAGATGGTATGATGCAATGCGG 771
DB 371 AGGCTAGTGTGATCTAGTCTTATATCTCCATATAGGAGACCGAGAGTCTTGTCG 430
QY 772 TGCTTACTTCCACATCTTAACCTTATGAAGTATTTATGATTTGCCCTTAAATATG 831
DB 431 TGCACCTGTTGCACAGGTAGTTTATTCAGCTTCTTGAACATGCTCTGGAATTGTG 490
QY 832 TGAAGCTGTGATCTTAAAGGCTATACAAAGCTTCGACGTCACAGGAAGATTAAAGTTT 891
DB 491 TGAAGCAAGGATCCAAAGGCTTTTACAAAGCTTCTGTCGCAAGAAA-AATAAAGGGTT 549
QY 892 CACTGGAATGATGATCCATACGAACCACTTAATGATGATGATG 938
DB 550 TACTGTGTGATGATCCATATGAAGCGCCATTGAATTGCGAGATG 596

RESULT 13
LOCUS BG126657 748 bp mRNA linear EST 31-JAN-2001
DEFINITION EST472303 tomato shoot/meristem Lycopersicon esculentum cDNA clone
cTOF13A21 5' sequence, mRNA sequence.

ACCESSION BG126657
VERSION BG126657.1 GI:12626845
KEYWORDS EST.
SOURCE tomato.

ORGANISM
Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.

REFERENCE
1 (bases 1 to 748)
van der Hoeven R., Bezzerides J., Sun H., Cho J., Utterback T.,
Hansen C., Rönning C. and Tanksley S.
Generation of ESTs from tomato shoot/meristem tissue
Unpublished (2001)
Contact: CUGI
Clemson University Genomics Institute
Clemson University

FEATURES
source

100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html.

Location/Qualifiers
1..748
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cTOF13A21"
/clone_lib="tomato shoot/meristem"
/tissue_type="shoot/meristem"
/dev_stage="developing shoots from 4-6wks old plants"
/lab_host="SOLR"
/note="Vector: pBluescript SK(-); Site1: EcoRI; Site2:
XhoI; Small expanding leaves from the growing tip were
taken from greenhouse plants (4-6wks old TA496). Tissue
was immediately frozen in liquid nitrogen."

BASE COUNT 227 a 129 c 176 g 216 t

Query Match 22.6%; Score 275.2; DB 10; Length 748;
Best Local Similarity 68.2%; Pred. No. 6.6e-31;
Matches 404; Conservative 0; Mismatches 173; Indels 15; Gaps 1;

QY 408 TAATGTCGAACATTCGGAAATCGACTAATATTTATGGCACAATTCCTTGATTGGACAAT 467
DB 120 TGATGCTACTGTGGCAATTCGGCAACATATCTGGCATGAAATCCAGTTGGGAAG 179
QY 468 CTGATAGACAGAAATTCGTGGGACAAAAGGCTGTGCTATGGATTAACAGACTCAGTG 527
DB 180 CGGAAAGGAAAAGCTGCTTAACCAACAGGAGTGTGTTGATGGATCAGGGTCTCAGT 239
QY 528 GTTCAGGAAAAGTACTCTTCGATGTCGACTGAGTTCGATTCGATTCGAGAGCCACC 587
DB 240 GATCAGAAAAGCACACTTCGATGTTCCCTAGTAGAGTTGCAGTCAAGGGTAAGC 299
QY 588 TCAGTATGATCTGATGTCGACACCTCAGACATGCGCTTAATAGATTTAAGCTTTA 647
DB 300 TTTTCATACGTTCTGATGTCGACACCTTCGGCATGGTCTGAAACAAGAAATCTTGGGTCT 359
QY 648 AGGCAGAAAGACCTTCGACAAAATATACGAAGTTCGTAAGTGGCAAGCTTTTTCGTG 707
DB 360 CACCAGAAAGCCGAACTAGAGATATACGTAGGACTGGGAAGTTGCAAAATCTCTTTGCG 419
QY 708 ATGCTGTGTCATATGATCTGATGATATCCATACAGGAGATCGTATGTCAT 767
DB 420 ATGCTGATTAATTTGCAFTTTCGATTTGATATCTCCTTACAGAAAGATCGTATGCTT 479
QY 768 GCCGTGCTCTACTTCCACATCTTAACCTTTATTTAAGTATTTATGATTTGCCCTTAAAA 827
DB 480 GCCGTGCTTTATTTGCCAGATAAAAGTTTCATTGAGGTTTTTATGAATATGCTCTACAAC 539
QY 828 TTTGTGAAGCTCGTATCTTAAGGCTTATACAGCTTCACGTACAGGAAGATTAA-- 885
DB 540 TGTGTGAAGATAGATACAAAGGCCCTCTACAAGCTAGCTCGTAGGGAATAAATCAAG 599
QY 886 -----AGGTTTCATCGAATTCGATTCATACGAAACCAACCAATTAATGCTG 932
DB 600 CACCTCATTTTGTAGGCTTTTACTGGAATAGATGATCCTTATGAACCACTTTGAATGCTG 659
QY 933 AGATAGTAATTAAGATGAAGATGAGGAATGCCCTTCACCAAGCAATGCC 984
DB 660 AGATTGAATACAACTTAAAGACGGAGTAGTTCTACACACATAGATGATG 711

RESULT 14
LOCUS BG120536 576 bp mRNA linear EST 29-NOV-2001
DEFINITION sah56h02.y1 Gm-cl049 Glycine max cDNA clone GENOME SYSTEMS CLONE
ID: Gm-cl049-2643 5' similar to SW:KAPS_CATRO_049204
ADENYLYLSULFATE KINASE, CHLOROPLAST PRECURSOR ;, mRNA sequence.
ACCESSION BG120536
VERSION BG120536.1 GI:14999722
KEYWORDS EST.

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 2, 2002, 00:28:47 ; Search time 50 Seconds
(without alignments)
761.966 Million cell updates/sec

Title: US-09-720-384A-4
Perfect score: 1826
Sequence: 1 RPFHFNTEPLVTHQTQP.....PKMAKQVLCYLENGYLQA 343

Scoring table: BLOSOM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802.*
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1826	100.0	343	21 AAY44789	Corn Adenylsulph
2	887	48.6	224	21 AAY44788	Corn Adenylsulph
3	858	47.0	246	21 AAY44792	Wheat Adenylsulph
4	852	46.7	311	21 AAY44791	Soybean Adenylsulph
5	834	45.7	208	21 AAG35788	Arabidopsis thalia
6	817	44.7	259	21 AAG14477	Arabidopsis thalia
7	817	44.7	272	21 AAG14476	Arabidopsis thalia
8	816	44.7	251	21 AAG11017	Arabidopsis thalia
9	816	44.7	251	21 AAG38348	Arabidopsis thalia
10	816	44.7	263	21 AAG11016	Arabidopsis thalia
11	816	44.7	263	21 AAG38347	Arabidopsis thalia

12	816	44.7	276	21 AAG11015	Arabidopsis thalia
13	816	44.7	276	21 AAG38346	Arabidopsis thalia
14	814	44.6	236	21 AAG14478	Arabidopsis thalia
15	809.5	44.3	237	21 AAG53120	Arabidopsis thalia
16	777.5	42.6	293	21 AAG29425	Arabidopsis thalia
17	777.5	42.6	293	21 AAY77957	A. thaliana enviro
18	757.5	41.5	252	21 AAG47458	Arabidopsis thalia
19	757.5	41.5	305	21 AAG47457	Arabidopsis thalia
20	757.5	41.5	310	21 AAG47456	Arabidopsis thalia
21	752.5	41.2	252	21 AAG14580	Arabidopsis thalia
22	752.5	41.2	305	21 AAG14579	Arabidopsis thalia
23	752.5	41.2	310	21 AAG14578	Arabidopsis thalia
24	617.5	33.8	161	21 AAG53456	Arabidopsis thalia
25	495	27.1	624	19 AAW70494	Human disease rela
26	495	27.1	624	20 AAY22349	Human APS kinase/A
27	495	27.1	625	20 AAW67882	Human secreted pro
28	492.5	27.0	648	22 AUA32438	Novel human secret
29	489	26.8	118	21 AAY44790	Rice Adenylsulph
30	481	26.3	201	22 AAG98409	Escherichia coli p
31	474	26.0	619	21 AAY79214	Human transferase
32	435.5	23.8	635	22 ABE71650	Drosophila melanog
33	390	21.4	139	21 AAG48931	Arabidopsis thalia
34	390	21.4	151	21 AAG48930	Arabidopsis thalia
35	390	21.4	164	21 AAG48929	Arabidopsis thalia
36	314.5	17.2	174	22 AAB96318	Putative P. abyssi
37	256.5	14.0	148	21 AAG12628	Zea mays protein f
38	249	13.6	58	21 AAY44793	Wheat Adenylsulph
39	202.5	11.1	65	22 ABB40266	Peptide #7772 enco
40	202.5	11.1	65	22 ABB24683	Protein #6682 enco
41	202.5	11.1	65	22 AAM61056	Human brain expres
42	202.5	11.1	65	22 AAM73760	Human bone marrow
43	202.5	11.1	65	22 AAM33947	Peptide #7984 enco
44	150.5	8.2	160	21 AAB32534	S. lavendulae MmCU
45	117	6.4	361	22 AAB94709	Human protein sequ

ALIGNMENTS

RESULT 1
AAY44789
ID AAY44789 standard; Protein; 343 AA.
XX
AC AAY44789;
XX
DT 04-MAY-2000 (first entry)
XX
XX
DE Corn Adenylsulphate Kinase-2.
XX
KW Adenylsulphate kinase; Adenosine-5'-phosphosulphate kinase; APS kinase;
KW 3'-Phospho-adenosine-5'-phosphosulphate; PAPS; sulphate assimilation;
KW corn; clone p0016.ctscj40rb; transgenic plant; screen; antibody.
XX
OS Zea mays.
XX
PN WO200004165-A1.
XX
PD 27-JAN-2000.
XX
PF 13-JUL-1999; 99WO-US15809.
XX
PR 14-JUL-1998; 98US-0092833.
XX
XX (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Falco SC, Allen SM, Anderson SL;
XX
DR WPI; 2000-182430/16.
XX
DR N-PSDB; AAZ50160.
XX
PT New nucleic acid molecule and chimeric gene encoding an adenosine-5'-
PT phosphosulphate kinase, useful for altering expression of sulfate
PT assimilation protein in plants -

PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134319.
PR 14-MAY-1999; 99US-0134421.
PR 18-MAY-1999; 99US-0134370.
PR 19-MAY-1999; 99US-0134768.
PR 20-MAY-1999; 99US-0134941.
PR 21-MAY-1999; 99US-0135124.
PR 24-MAY-1999; 99US-0135353.
PR 25-MAY-1999; 99US-0135629.
PR 27-MAY-1999; 99US-0136021.
PR 28-MAY-1999; 99US-0136392.
PR 01-JUN-1999; 99US-0136782.
PR 03-JUN-1999; 99US-0137222.
PR 04-JUN-1999; 99US-0137528.
PR 07-JUN-1999; 99US-0137502.
PR 08-JUN-1999; 99US-0137724.
PR 10-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 22-JUN-1999; 99US-0139817.
PR 23-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 28-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141482.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142920.
PR 13-JUL-1999; 99US-0142977.
PR 14-JUL-1999; 99US-0143542.
PR 15-JUL-1999; 99US-0143624.
PR 99US-0144005.

PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149923.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.

PR	07-OCT-1999;	99US-0158029.	PF	25-FEB-2000;	2000EP-0301439.
PR	08-OCT-1999;	99US-0158232.	XX	25-FEB-1999;	99US-0121825.
PR	12-OCT-1999;	99US-0158369.	PR	05-MAR-1999;	99US-0123180.
PR	13-OCT-1999;	99US-0159293.	PR	09-MAR-1999;	99US-0123548.
PR	13-OCT-1999;	99US-0159294.	PR	23-MAR-1999;	99US-0125788.
PR	13-OCT-1999;	99US-0159295.	PR	25-MAR-1999;	99US-0126264.
PR	14-OCT-1999;	99US-0159329.	PR	29-MAR-1999;	99US-0126785.
PR	14-OCT-1999;	99US-0159330.	PR	01-APR-1999;	99US-0127462.
PR	14-OCT-1999;	99US-0159331.	PR	06-APR-1999;	99US-0128234.
PR	14-OCT-1999;	99US-0159637.	PR	08-APR-1999;	99US-0128714.
PR	14-OCT-1999;	99US-0159638.	PR	16-APR-1999;	99US-0129845.
PR	21-OCT-1999;	99US-0159584.	PR	19-APR-1999;	99US-0130077.
PR	21-OCT-1999;	99US-0160741.	PR	21-APR-1999;	99US-0130449.
PR	21-OCT-1999;	99US-0160767.	PR	23-APR-1999;	99US-0130510.
PR	21-OCT-1999;	99US-0160788.	PR	23-APR-1999;	99US-0130891.
PR	21-OCT-1999;	99US-0160770.	PR	28-APR-1999;	99US-0131449.
PR	21-OCT-1999;	99US-0160814.	PR	30-APR-1999;	99US-0132048.
PR	22-OCT-1999;	99US-0160815.	PR	30-APR-1999;	99US-0132407.
PR	22-OCT-1999;	99US-0160980.	PR	04-MAY-1999;	99US-0132484.
PR	22-OCT-1999;	99US-0160981.	PR	05-MAY-1999;	99US-0132485.
PR	25-OCT-1999;	99US-0160989.	PR	06-MAY-1999;	99US-0132486.
PR	25-OCT-1999;	99US-0161404.	PR	06-MAY-1999;	99US-0132487.
PR	25-OCT-1999;	99US-0161405.	PR	07-MAY-1999;	99US-0132863.
PR	25-OCT-1999;	99US-0161406.	PR	11-MAY-1999;	99US-0134256.
PR	26-OCT-1999;	99US-0161359.	PR	14-MAY-1999;	99US-0134218.
PR	26-OCT-1999;	99US-0161360.	PR	14-MAY-1999;	99US-0134219.
PR	26-OCT-1999;	99US-0161361.	PR	14-MAY-1999;	99US-0134221.
PR	28-OCT-1999;	99US-0161920.	PR	14-MAY-1999;	99US-0134370.
PR	28-OCT-1999;	99US-0161932.	PR	18-MAY-1999;	99US-0134768.
PR	28-OCT-1999;	99US-0161993.	PR	19-MAY-1999;	99US-0134941.
PR	29-OCT-1999;	99US-0162142.	PR	20-MAY-1999;	99US-0135124.
Query Match 45 78; Score 834; DB 21; Length 208;					99US-0135353.
Best Local Similarity 74.88; Pred. No. 4,3e-74;					99US-0135529.
Matches 154; Conservative 26; Mismatches 24; Indels 0; Gaps 0;					99US-0136021.
QY	137	MSNIGKSTNTLHNCILGSDRQKLLGQKGVVWITGLSGSKSTLACLSRELHCRGHL 196	PR	27-MAY-1999;	99US-0136392.
DB	1	MSTVGNSTNFMQESTFQKLLNQKGVVWITGLSGSKSTLACLSRELNNRGL 60	PR	28-MAY-1999;	99US-0136782.
QY	197	TYVLGDNLRHGLNRDLKFKAEDRAENIRRVGEVAKLFADAGVICIASLISPYRRDRDAC 256	PR	01-JUN-1999;	99US-0137222.
DB	61	SYLLDGNLRHGLNRDLKFKAEDRAENIRRVGEVAKLFADAGVICIASLISPYEKDRDAC 120	PR	03-JUN-1999;	99US-0137528.
QY	257	RALLPHNFTEVFIDPLKICEARDPKGLYKLARTGKIKGFTGIDDPYEPPINGEIVIKM 316	PR	04-JUN-1999;	99US-0137502.
DB	121	REMIQNSFTFEVPMNMSQLCEARDPKGLYKLARAGKIKGFTGIDDPYESPLNCETELKE 180	PR	07-JUN-1999;	99US-0137724.
QY	317	KDEECPSPKAMAKQVLCYLENGYLQ 342	PR	08-JUN-1999;	99US-0138094.
DB	181	KEGECPSPVMAEEVISYLEDKGFLO 206	PR	10-JUN-1999;	99US-0138540.
RESULT 6					99US-0138847.
AAG14477					99US-0139119.
ID	AAG14477 standard; Protein; 259 AA.				99US-0139452.
XX	AAG14477;				99US-0139453.
AC	AAG14477;				99US-0139492.
XX	17-OCT-2000 (first entry)				99US-0139454.
DT	Arabidopsis thaliana protein fragment SEQ ID NO: 14355.				99US-0139455.
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 14355.				99US-0139456.
XX	Protein identification; signal transduction pathway; metabolic pathway;				99US-0139457.
KW	hybridisation assay; genetic mapping; gene expression control; promoter;				99US-0139458.
KW	termination sequence.				99US-0139459.
XX	Arabidopsis thaliana.				99US-0139460.
OS	Arabidopsis thaliana.				99US-0139461.
XX	EP1033405-A2.				99US-0139462.
PN	06-SEP-2000.				99US-0139463.
XX					99US-0139750.
XX					99US-0139763.
XX					99US-0139817.
XX					99US-0139899.
XX					99US-0140353.
XX					99US-0140354.
XX					99US-0140695.
XX					99US-0140823.
XX					99US-0140991.
XX					99US-0141287.
XX					99US-0141842.
XX					99US-0142154.
XX					99US-0142055.
XX					99US-0142390.
XX					99US-0142803.

PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145313.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147360.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.

PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162143.

Query Match 44.7%; Score 817; DB 21; Length 259;
Best Local Similarity 60.6%; Pred. No. 2.8e-72;
Matches 163; Conservative 32; Mismatches 52; Indels 22; Gaps 4;
QY 87 LADSGRGEGR-----GARTHCRGIGRWVRRRRNCAPGPHSPVKEK 134
DB 1 MASPGLSDNSRSVVVACVSMDSQTLSHNKNKSIPELKSIING-----HTGQKQG 53
QY 135 PVMSNICKSTNLIWHNCLIGOSDROKLLGQKGVVWITGLSGSGKSTLACALSRELHCRG 194
DB 54 P-LSTVGNSTNIKWHCEKSVKVDRLDQKGVVWITGLSGSGKSTLACALNQLYQKG 112
QY 195 HLTVVLDGDNLRHGLNRDLSFKAEDRAENIRRVGEVAKLPADAGVICTIASLISPYRRDRD 254
DB 113 KLCYILDGDNVNRHGLNRDLSFKAEDRAENIRRVGEVAKLPADAGIICIASLISPYRTDRD 172
QY 255 ACRALLPHSNFIEVFIDLPLKICEARDPKGLYKLARTGKIKGTGIDDPYEPPIGEIIV 314
DB 173 ACRNLLPEGDFVEFMDVPLSVCEARDPKGLYKLARAGIKGTGIDDPYEPPLNCE--I 230
QY 315 KMKDECPSPKAMAKQVLCYLEENGYLQA 343
DB 231 SLGREGGTSPTEMAEKVGVGLDNKGYLQA 259

RESULT 7
AAG14476
ID AAG14476 standard; Protein; 272 AA.
XX
AC AAG14476;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 14354.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;

KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
XX EP1033405-A2.
PN
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.

PR	07-SEP-1999;	99US-0152363.		ID	AG11017 standard; Protein; 251 AA.
PR	10-SEP-1999;	99US-0153070.		XX	
PR	13-SEP-1999;	99US-0153758.		AC	AAG11017;
PR	16-SEP-1999;	99US-0154018.		XX	
PR	20-SEP-1999;	99US-0154039.		DT	17-OCT-2000 (first entry)
PR	22-SEP-1999;	99US-0154779.		XX	
PR	23-SEP-1999;	99US-0155139.		DE	Arabidopsis thaliana protein fragment SEQ ID NO: 9573.
PR	24-SEP-1999;	99US-0155486.		XX	
PR	28-SEP-1999;	99US-0155659.		KW	Protein identification; signal transduction pathway; metabolic pathway;
PR	28-SEP-1999;	99US-0156458.		KW	hybridisation assay; genetic mapping; gene expression control; promoter;
PR	29-SEP-1999;	99US-0156596.		KW	termination sequence.
PR	04-OCT-1999;	99US-0157117.		XX	
PR	05-OCT-1999;	99US-0157753.		OS	Arabidopsis thaliana.
PR	06-OCT-1999;	99US-0157865.		XX	
PR	07-OCT-1999;	99US-0158029.		PN	EP1033405-A2.
PR	08-OCT-1999;	99US-0158232.		XX	
PR	12-OCT-1999;	99US-0158369.		PD	06-SEP-2000.
PR	13-OCT-1999;	99US-0159293.		XX	
PR	13-OCT-1999;	99US-0159294.		PF	25-FEB-2000; 2000EP-0301439.
PR	13-OCT-1999;	99US-0159295.		XX	
PR	14-OCT-1999;	99US-0159329.		PR	25-FEB-1999; 99US-0121825.
PR	14-OCT-1999;	99US-0159330.		PR	05-MAR-1999; 99US-0123180.
PR	14-OCT-1999;	99US-0159331.		PR	09-MAR-1999; 99US-0123548.
PR	14-OCT-1999;	99US-0159637.		PR	23-MAR-1999; 99US-0125788.
PR	14-OCT-1999;	99US-0159638.		PR	25-MAR-1999; 99US-0126264.
PR	18-OCT-1999;	99US-0159584.		PR	29-MAR-1999; 99US-0126785.
PR	21-OCT-1999;	99US-0160741.		PR	01-APR-1999; 99US-0127462.
PR	21-OCT-1999;	99US-0160767.		PR	06-APR-1999; 99US-0128234.
PR	21-OCT-1999;	99US-0160768.		PR	08-APR-1999; 99US-0128714.
PR	21-OCT-1999;	99US-0160770.		PR	16-APR-1999; 99US-0129845.
PR	21-OCT-1999;	99US-0160814.		PR	19-APR-1999; 99US-0130077.
PR	21-OCT-1999;	99US-0160815.		PR	21-APR-1999; 99US-0130449.
PR	22-OCT-1999;	99US-0160980.		PR	23-APR-1999; 99US-0130510.
PR	22-OCT-1999;	99US-0160981.		PR	23-APR-1999; 99US-0130891.
PR	22-OCT-1999;	99US-0160989.		PR	28-APR-1999; 99US-0131449.
PR	25-OCT-1999;	99US-0161404.		PR	30-APR-1999; 99US-0132048.
PR	25-OCT-1999;	99US-0161405.		PR	30-APR-1999; 99US-0132407.
PR	25-OCT-1999;	99US-0161406.		PR	04-MAY-1999; 99US-0132484.
PR	26-OCT-1999;	99US-0161359.		PR	05-MAY-1999; 99US-0132485.
PR	26-OCT-1999;	99US-0161360.		PR	06-MAY-1999; 99US-0132486.
PR	26-OCT-1999;	99US-0161361.		PR	06-MAY-1999; 99US-0132487.
PR	28-OCT-1999;	99US-0161920.		PR	07-MAY-1999; 99US-0132863.
PR	28-OCT-1999;	99US-0161992.		PR	11-MAY-1999; 99US-0134256.
PR	28-OCT-1999;	99US-0161993.		PR	14-MAY-1999; 99US-0134218.
PR	28-OCT-1999;	99US-0161993.		PR	14-MAY-1999; 99US-0134219.
PR	29-OCT-1999;	99US-0162142.		PR	14-MAY-1999; 99US-0134221.
PR				PR	14-MAY-1999; 99US-0134370.
PR				PR	18-MAY-1999; 99US-0134768.
PR				PR	19-MAY-1999; 99US-0134941.
PR				PR	20-MAY-1999; 99US-0135124.
PR				PR	21-MAY-1999; 99US-0135353.
PR				PR	24-MAY-1999; 99US-0135629.
PR				PR	25-MAY-1999; 99US-0136021.
PR				PR	27-MAY-1999; 99US-0136392.
PR				PR	28-MAY-1999; 99US-0136782.
PR				PR	01-JUN-1999; 99US-0137222.
PR				PR	03-JUN-1999; 99US-0137528.
PR				PR	04-JUN-1999; 99US-0137502.
PR				PR	07-JUN-1999; 99US-0137724.
PR				PR	08-JUN-1999; 99US-0138094.
PR				PR	10-JUN-1999; 99US-0138540.
PR				PR	10-JUN-1999; 99US-0138847.
PR				PR	14-JUN-1999; 99US-0139119.
PR				PR	16-JUN-1999; 99US-0139452.
PR				PR	16-JUN-1999; 99US-0139453.
PR				PR	17-JUN-1999; 99US-0139492.
PR				PR	18-JUN-1999; 99US-0139454.
PR				PR	18-JUN-1999; 99US-0139455.
PR				PR	18-JUN-1999; 99US-0139456.
PR				PR	18-JUN-1999; 99US-0139457.
PR				PR	18-JUN-1999; 99US-0139458.
PR				PR	18-JUN-1999; 99US-0139459.
PR				PR	18-JUN-1999; 99US-0139460.
Query Match 44.7%; Score 817; DB 21; Length 272;					
Best Local Similarity 60.8%; Pred. No. 3.1e-72;					
Matches 163; Conservative 32; Mismatches 52; Indels 22; Gaps 4;					
Qy	87	LASDSGRREGGR-----GARTHCHRGIGRWVRRRRRNGAAGPEAPHSVPVKEK 134			
Db	14	MASPKGLSDSNSRSVVVYVACVSMDSQTLSHKNGSIPELKLSNG-----HTGQKQG 66			
Qy	135	PWMSNIGKSTNLWHNCLIQSDROKLLGQKGVVWITGLSGSGKSTLACALSRELHCRG 194			
Db	67	P-LSTVGNSTNKKHESVBEKVDQRLLDQKGVITWVTLGSGSGKSTLACALNQMLYQKG 125			
Qy	195	HTYYLDGDNLRHGLNRLSFKAEADRAENIRRVGEVAKLFADAGVICIASLISPYRRDRD 254			
Db	126	KLCYILDGDNVHRHGLNRLSFKAEADRAENIRRVGEVAKLFADAGIICIASLISPYRTDRD 185			
Qy	255	ACRALLPHSNTEVFDLPKLGCEARDPKGLYKARTCKIKGFTGIDDPYEPPIGEIIV 314			
Db	186	ACRNLLPEGDVFEVMDVPLVSCVARDPKGLYKLARAGIKGFTGIDDPYEPPLNCE--I 243			
Qy	315	KMKDECPSPKAMAKQVLCYLEENGYLQA 343			
Db	244	SLREGGTSPIEMAEKVGVLDNKGYLQA 272			
RESULT 8					
AAG11017					

PR	18-JUN-1999;	99US-0139461.	23-AUG-1999;	99US-0149930.	
PR	18-JUN-1999;	99US-0139462.	25-AUG-1999;	99US-0150566.	
PR	18-JUN-1999;	99US-0139463.	26-AUG-1999;	99US-0150884.	
PR	18-JUN-1999;	99US-0139750.	27-AUG-1999;	99US-0151065.	
PR	18-JUN-1999;	99US-0139763.	27-AUG-1999;	99US-0151066.	
PR	21-JUN-1999;	99US-0139817.	27-AUG-1999;	99US-0151080.	
PR	22-JUN-1999;	99US-0139899.	PR	99US-0151303.	
PR	23-JUN-1999;	99US-0140353.	PR	99US-0151303.	
PR	23-JUN-1999;	99US-0140354.	PR	99US-0151438.	
PR	24-JUN-1999;	99US-0140695.	PR	99US-0151930.	
PR	28-JUN-1999;	99US-0140823.	PR	99US-0152363.	
PR	29-JUN-1999;	99US-0140991.	PR	99US-0153070.	
PR	30-JUN-1999;	99US-0141287.	PR	99US-0153758.	
PR	01-JUL-1999;	99US-0141842.	PR	99US-0154018.	
PR	02-JUL-1999;	99US-0142154.	PR	99US-0154039.	
PR	06-JUL-1999;	99US-0142055.	PR	99US-0154779.	
PR	08-JUL-1999;	99US-0142390.	PR	99US-0155139.	
PR	08-JUL-1999;	99US-0142803.	PR	99US-0155486.	
PR	09-JUL-1999;	99US-0142320.	PR	99US-0155659.	
PR	12-JUL-1999;	99US-0142977.	PR	99US-0156458.	
PR	13-JUL-1999;	99US-0143542.	PR	99US-0156596.	
PR	14-JUL-1999;	99US-0143624.	PR	99US-0156596.	
PR	15-JUL-1999;	99US-0144005.	PR	99US-0157117.	
PR	16-JUL-1999;	99US-0144085.	PR	99US-0157753.	
PR	16-JUL-1999;	99US-0144085.	PR	99US-0157865.	
PR	19-JUL-1999;	99US-0144325.	PR	99US-0158029.	
PR	19-JUL-1999;	99US-0144325.	PR	99US-0158232.	
PR	19-JUL-1999;	99US-0144331.	PR	99US-0158369.	
PR	19-JUL-1999;	99US-0144332.	PR	99US-0159293.	
PR	19-JUL-1999;	99US-0144333.	PR	99US-0159294.	
PR	19-JUL-1999;	99US-0144334.	PR	99US-0159295.	
PR	19-JUL-1999;	99US-0144335.	PR	99US-0159329.	
PR	20-JUL-1999;	99US-0144352.	PR	99US-0159330.	
PR	20-JUL-1999;	99US-0144532.	PR	99US-0159330.	
PR	20-JUL-1999;	99US-0144884.	PR	99US-0159331.	
PR	21-JUL-1999;	99US-0144814.	PR	99US-0159637.	
PR	21-JUL-1999;	99US-0145086.	PR	99US-0159638.	
PR	21-JUL-1999;	99US-0145088.	PR	99US-0159584.	
PR	22-JUL-1999;	99US-0145085.	PR	99US-0160741.	
PR	22-JUL-1999;	99US-0145087.	PR	99US-0160767.	
PR	22-JUL-1999;	99US-0145089.	PR	99US-0160768.	
PR	22-JUL-1999;	99US-0145192.	PR	99US-0160770.	
PR	23-JUL-1999;	99US-0145145.	PR	99US-0160814.	
PR	23-JUL-1999;	99US-0145218.	PR	99US-0160815.	
PR	23-JUL-1999;	99US-0145224.	PR	99US-0160980.	
PR	26-JUL-1999;	99US-0145276.	PR	99US-0160981.	
PR	27-JUL-1999;	99US-0145913.	PR	99US-0160989.	
PR	27-JUL-1999;	99US-0145918.	PR	99US-0161404.	
PR	27-JUL-1999;	99US-0145919.	PR	99US-0161405.	
PR	28-JUL-1999;	99US-0145951.	PR	99US-0161406.	
PR	02-AUG-1999;	99US-0146386.	PR	99US-0161359.	
PR	02-AUG-1999;	99US-0146388.	PR	99US-0161360.	
PR	03-AUG-1999;	99US-0146389.	PR	99US-0161361.	
PR	03-AUG-1999;	99US-0147038.	PR	99US-0161920.	
PR	04-AUG-1999;	99US-0147204.	PR	99US-0161992.	
PR	04-AUG-1999;	99US-0147302.	PR	99US-0161993.	
PR	05-AUG-1999;	99US-0147192.	PR	99US-0162142.	
PR	05-AUG-1999;	99US-0147260.			
PR	06-AUG-1999;	99US-0147303.			
PR	06-AUG-1999;	99US-0147416.			
PR	09-AUG-1999;	99US-0147493.			
PR	09-AUG-1999;	99US-0147935.			
PR	10-AUG-1999;	99US-0148171.			
PR	11-AUG-1999;	99US-0148319.			
PR	12-AUG-1999;	99US-0148341.			
PR	13-AUG-1999;	99US-0148565.			
PR	13-AUG-1999;	99US-0148684.			
PR	16-AUG-1999;	99US-0149368.			
PR	17-AUG-1999;	99US-0149175.			
PR	18-AUG-1999;	99US-0149426.			
PR	20-AUG-1999;	99US-0149722.			
PR	20-AUG-1999;	99US-0149723.			
PR	20-AUG-1999;	99US-0149929.			
PR	23-AUG-1999;	99US-0149902.			

PR	23-AUG-1999;	99US-0149930.	23-AUG-1999;	99US-0149930.	
PR	25-AUG-1999;	99US-0150566.	25-AUG-1999;	99US-0150566.	
PR	26-AUG-1999;	99US-0150884.	26-AUG-1999;	99US-0150884.	
PR	27-AUG-1999;	99US-0151065.	27-AUG-1999;	99US-0151065.	
PR	27-AUG-1999;	99US-0151066.	27-AUG-1999;	99US-0151066.	
PR	27-AUG-1999;	99US-0151080.	27-AUG-1999;	99US-0151080.	
PR	30-AUG-1999;	99US-0151303.	30-AUG-1999;	99US-0151303.	
PR	31-AUG-1999;	99US-0151438.	31-AUG-1999;	99US-0151438.	
PR	01-SEP-1999;	99US-0151930.	01-SEP-1999;	99US-0151930.	
PR	07-SEP-1999;	99US-0152363.	07-SEP-1999;	99US-0152363.	
PR	10-SEP-1999;	99US-0153070.	10-SEP-1999;	99US-0153070.	
PR	13-SEP-1999;	99US-0153758.	13-SEP-1999;	99US-0153758.	
PR	15-SEP-1999;	99US-0154018.	15-SEP-1999;	99US-0154018.	
PR	16-SEP-1999;	99US-0154039.	16-SEP-1999;	99US-0154039.	
PR	20-SEP-1999;	99US-0154779.	20-SEP-1999;	99US-0154779.	
PR	22-SEP-1999;	99US-0155139.	22-SEP-1999;	99US-0155139.	
PR	23-SEP-1999;	99US-0155486.	23-SEP-1999;	99US-0155486.	
PR	24-SEP-1999;	99US-0155659.	24-SEP-1999;	99US-0155659.	
PR	28-SEP-1999;	99US-0156458.	28-SEP-1999;	99US-0156458.	
PR	29-SEP-1999;	99US-0156596.	29-SEP-1999;	99US-0156596.	
PR	04-OCT-1999;	99US-0157117.	04-OCT-1999;	99US-0157117.	
PR	05-OCT-1999;	99US-0157753.	05-OCT-1999;	99US-0157753.	
PR	06-OCT-1999;	99US-0157865.	06-OCT-1999;	99US-0157865.	
PR	07-OCT-1999;	99US-0158029.	07-OCT-1999;	99US-0158029.	
PR	08-OCT-1999;	99US-0158232.	08-OCT-1999;	99US-0158232.	
PR	12-OCT-1999;	99US-0158369.	12-OCT-1999;	99US-0158369.	
PR	13-OCT-1999;	99US-0159293.	13-OCT-1999;	99US-0159293.	
PR	13-OCT-1999;	99US-0159294.	13-OCT-1999;	99US-0159294.	
PR	13-OCT-1999;	99US-0159329.	13-OCT-1999;	99US-0159329.	
PR	14-OCT-1999;	99US-0159330.	14-OCT-1999;	99US-0159330.	
PR	14-OCT-1999;	99US-0159331.	14-OCT-1999;	99US-0159331.	
PR	14-OCT-1999;	99US-0159637.	14-OCT-1999;	99US-0159637.	
PR	14-OCT-1999;	99US-0159638.	14-OCT-1999;	99US-0159638.	
PR	18-OCT-1999;	99US-0159584.	18-OCT-1999;	99US-0159584.	
PR	21-OCT-1999;	99US-0160741.	21-OCT-1999;	99US-0160741.	
PR	21-OCT-1999;	99US-0160767.	21-OCT-1999;	99US-0160767.	
PR	21-OCT-1999;	99US-0160768.	21-OCT-1999;	99US-0160768.	
PR	21-OCT-1999;	99US-0160770.	21-OCT-1999;	99US-0160770.	
PR	21-OCT-1999;	99US-0160814.	21-OCT-1999;	99US-0160814.	
PR	21-OCT-1999;	99US-0160815.	21-OCT-1999;	99US-0160815.	
PR	22-OCT-1999;	99US-0160980.	22-OCT-1999;	99US-0160980.	
PR	22-OCT-1999;	99US-0160981.	22-OCT-1999;	99US-0160981.	
PR	22-OCT-1999;	99US-0160989.	22-OCT-1999;	99US-0160989.	
PR	25-OCT-1999;	99US-0161404.	25-OCT-1999;	99US-0161404.	
PR	25-OCT-1999;	99US-0161405.	25-OCT-1999;	99US-0161405.	
PR	26-OCT-1999;	99US-0161406.	26-OCT-1999;	99US-0161406.	
PR	26-OCT-1999;	99US-0161359.	26-OCT-1999;	99US-0161359.	
PR	26-OCT-1999;	99US-0161360.	26-OCT-1999;	99US-0161360.	
PR	28-OCT-1999;	99US-0161361.	28-OCT-1999;	99US-0161361.	
PR	28-OCT-1999;	99US-0161920.	28-OCT-1999;	99US-0161920.	
PR	28-OCT-1999;	99US-0161992.	28-OCT-1999;	99US-0161992.	
PR	28-OCT-1999;	99US-0161993.	28-OCT-1999;	99US-0161993.	
PR	29-OCT-1999;	99US-0162142.	29-OCT-1999;	99US-0162142.	

Query Match

Best Local Similarity

Matches 159; Conservative

Score 816; DB 21; Pred. No. 3.4e-72;

Length 251;

QY	100	GARTCHRGIGRWVRRRRRNGAAPGEAP	---HSPYKEKPVMSNTGKSTNTLHNCILQGS	156
		:	: :	
Db	18	GSQTLSH	-----NKGSIPEVKSINGHTGQKQGP-LSTVGNSTNIKWHECSEVKV	66

QY	157	DRQKLLGQKGVVWITGLSGSGKSTLACALSRHLCRCHLHVYLDGDNLRHGLNRDLSPK	216
		: :	

Db 187 CEARDPKGLYKLARAGKIKGFTGIDDPYEPPLNCE--ISLREGGTSPIEMAEKVVGILD 244
QY 337 ENGYLQA 343
Db 245 NKGYLQA 251
RESULT 9
AAG38348
ID AAG38348 standard; Protein; 251 AA.
XX
AC AAG38348;
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 47297.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 23-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.

PR 01-JUN-1999; 99US-0137322.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0143977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 08-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match				44.7%;	Score 816;	DB 21;	Length 263;
Best Local Similarity				64.4%;	Pred. No. 3.7e-72;		
Matches 159;				Conservative 32;	Mismatches 40;	Indels 16;	Gaps 4;
QY	100	GARTHCRCIGRWRRRRRCAGP	---HSPVKRPVMSNIGKSTNILWHNCLIGOS	156			
DB	30	GSQTLSH-----	NKNGSIPVKSTINGTGQKQP-LSTVGNSTINIKWHECSVEKV	78			
QY	157	DROKLLGQKGVWITGLSSGKSTLACALSRHLHCRGHLTYVLDGDLNRHGLNRDLSPK	216				
DB	79	DRORLLDQKGVWITGLSSGKSTLACALNQMLYQKGLCYILDGDNVRHGLNRDLSPK	138				
QY	217	AEDRAENIRRVGEVAKLFAADGVICIASLISPYRRDRDRCALLPHSNFIEVFDLPLKI	276				
DB	139	AEDRAENIRRVGEVAKLFAADGVICIASLISPYRTDRDRCALLPEGDFVEVFMVPLSV	198				
QY	277	CEARDPKGLYKLIARTGKIGFTGDDPYEPPINGELIVKMKDECPSPKAMAKQVLCYLE	336				
DB	199	CEARDPKGLYKLIARAGKIGFTGDDPYEPLNCE--ISLREGGTSPIEMAEKVVGYL	256				
QY	337	ENGYLQA 343					
DB	257	NKGYLQA 263					
RESULT 11							
ID	AAG38347	standard; Protein; 263 AA.					
XX	AC	AAG38347;					
XX	AC	18-OCT-2000 (first entry)					
DT	XX	Arabidopsis thaliana protein fragment SEQ ID NO: 47296.					
DE	XX	Protein identification; signal transduction pathway; metabolic pathway;					
KW	XX	hybridisation assay; genetic mapping; gene expression control; promoter;					
KW	XX	termination sequence.					
XX	XX	Arabidopsis thaliana.					
XX	XX	EPI033405-A2.					
XX	XX	06-SEP-2000.					
XX	XX	25-FEB-2000; 2000EP-0301439.					
XX	XX	25-FEB-1999; 99US-0121825.					
PR	05-MAR-1999;	99US-0123180.					
PR	09-MAR-1999;	99US-0123548.					
PR	23-MAR-1999;	99US-0125788.					
PR	25-MAR-1999;	99US-0126264.					
PR	29-MAR-1999;	99US-0126785.					
PR	01-APR-1999;	99US-0127462.					
PR	06-APR-1999;	99US-0128234.					
PR	08-APR-1999;	99US-0128714.					
PR	16-APR-1999;	99US-0129845.					
PR	19-APR-1999;	99US-0130077.					
PR	21-APR-1999;	99US-0130449.					
PR	23-APR-1999;	99US-0130510.					
PR	28-APR-1999;	99US-0130891.					
PR	30-APR-1999;	99US-0131449.					
PR	30-APR-1999;	99US-0132048.					
PR	04-MAY-1999;	99US-0132407.					
PR	05-MAY-1999;	99US-0132484.					
PR	06-MAY-1999;	99US-0132485.					
PR	06-MAY-1999;	99US-0132486.					
PR	06-MAY-1999;	99US-0132487.					
PR	07-MAY-1999;	99US-0132863.					
PR	11-MAY-1999;	99US-0134256.					
PR	14-MAY-1999;	99US-0134218.					
PR	14-MAY-1999;	99US-0134219.					
PR	14-MAY-1999;	99US-0134221.					
PR	14-MAY-1999;	99US-0134370.					
PR	18-MAY-1999;	99US-0134768.					
PR	19-MAY-1999;	99US-0134941.					
PR	20-MAY-1999;	99US-0135124.					
PR	21-MAY-1999;	99US-0135353.					
PR	24-MAY-1999;	99US-0135629.					
PR	25-MAY-1999;	99US-0136021.					
PR	27-MAY-1999;	99US-0136392.					
PR	28-MAY-1999;	99US-0136782.					
PR	01-JUN-1999;	99US-0137222.					
PR	03-JUN-1999;	99US-0137528.					
PR	04-JUN-1999;	99US-0137502.					
PR	07-JUN-1999;	99US-0137724.					
PR	08-JUN-1999;	99US-0138094.					
PR	10-JUN-1999;	99US-0138540.					
PR	10-JUN-1999;	99US-0138847.					
PR	14-JUN-1999;	99US-0139119.					
PR	16-JUN-1999;	99US-0139452.					
PR	16-JUN-1999;	99US-0139453.					
PR	17-JUN-1999;	99US-0139492.					
PR	18-JUN-1999;	99US-0139454.					
PR	18-JUN-1999;	99US-0139455.					
PR	18-JUN-1999;	99US-0139456.					
PR	18-JUN-1999;	99US-0139457.					
PR	18-JUN-1999;	99US-0139458.					
PR	18-JUN-1999;	99US-0139459.					
PR	18-JUN-1999;	99US-0139460.					
PR	18-JUN-1999;	99US-0139461.					
PR	18-JUN-1999;	99US-0139462.					
PR	18-JUN-1999;	99US-0139463.					
PR	18-JUN-1999;	99US-0139750.					
PR	18-JUN-1999;	99US-0139763.					
PR	21-JUN-1999;	99US-0139817.					
PR	22-JUN-1999;	99US-0139899.					
PR	23-JUN-1999;	99US-0140353.					
PR	23-JUN-1999;	99US-0140354.					
PR	24-JUN-1999;	99US-0140695.					
PR	28-JUN-1999;	99US-0140823.					
PR	29-JUN-1999;	99US-0140991.					
PR	30-JUN-1999;	99US-0141287.					
PR	01-JUL-1999;	99US-0141842.					
PR	01-JUL-1999;	99US-0142154.					
PR	02-JUL-1999;	99US-0142055.					
PR	06-JUL-1999;	99US-0142390.					
PR	08-JUL-1999;	99US-0142803.					
PR	09-JUL-1999;	99US-0142920.					
PR	12-JUL-1999;	99US-0142977.					
PR	13-JUL-1999;	99US-0143542.					
PR	14-JUL-1999;	99US-0143624.					
PR	15-JUL-1999;	99US-0144005.					
PR	16-JUL-1999;	99US-0144085.					
PR	16-JUL-1999;	99US-0144086.					
PR	19-JUL-1999;	99US-0144325.					
PR	19-JUL-1999;	99US-0144331.					
PR	19-JUL-1999;	99US-0144332.					
PR	19-JUL-1999;	99US-0144333.					
PR	19-JUL-1999;	99US-0144334.					
PR	19-JUL-1999;	99US-0144335.					
PR	19-JUL-1999;	99US-0144352.					
PR	20-JUL-1999;	99US-0144632.					
PR	20-JUL-1999;	99US-0144884.					
PR	21-JUL-1999;	99US-0144814.					
PR	21-JUL-1999;	99US-0145086.					
PR	21-JUL-1999;	99US-0145088.					
PR	22-JUL-1999;	99US-0145085.					
PR	22-JUL-1999;	99US-0145087.					
PR	22-JUL-1999;	99US-0145089.					
PR	22-JUL-1999;	99US-0145192.					
PR	23-JUL-1999;	99US-0145145.					
PR	23-JUL-1999;	99US-0145218.					
PR	23-JUL-1999;	99US-0145224.					
PR	26-JUL-1999;	99US-0145276.					
PR	27-JUL-1999;	99US-0145913.					

```
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.

PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 44.7%; Score 816; DB 21; Length 263;
Best Local Similarity 64.4%; Pred. No. 3.7e-72;
Matches 159; Conservative 32; Mismatches 40; Indels 16; Gaps 4;

QY 100 GARTHCHRGIGRWRRRRRNGAAGEAP---HSPVKEKPMNSNIGKSTNLIWHNCLIGQS 156
      |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
DB 30 GSOTLSH-----NKGSIPEVKSINGHTGQKGP-LSTVGNSTNIIKHECSVEKV 78
      |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|

QY 157 DRKLLGQKGCYVWITGLSGSKSTLACALSRELHCRGHLTVLDGDNLRHGLNRDLSFK 216
      |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
DB 79 DRQELLDQKGCYIWTGLSGSKSTLACALNQMLYQKGLCYILDGDNVRHGLNRDLSFK 138
      |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|

QY 217 AEDRAENIRRVGEVAKLFADAGVICIASLSPYRRDRDACRALLPHISNFIEVFDLPKI 276
      |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
DB 139 AEDRAENIRRVGEVAKLFADAGIIICIASLSPYRTDRDACRSLPGLPEGDFVEFMDVPLSV 198
      |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|

QY 277 CEARDPKGLYKLARTGKIKGTGIDDDPYEPPINGEIVIKMKDECECPKAMAKOVLCYLE 336
      |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
DB 199 CEARDPKGLYKLARAGKIKGTGIDDDPYEPPLNCE--ISLGREGGTSPIEMAEKVVGYLD 256
      |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|

QY 337 ENGYLQA 343
      |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
DB 257 NKGYLQA 263
      |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|

RESULT 12
AAG11015
ID AAG11015 standard; Protein; 276 AA.
XX
AC AAG11015;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 9571.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
```


PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 08-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 18-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.

XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134321.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151348.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 2, 2002, 01:49:58 : Search time 19 Seconds
(without alignments) updates/sec
440.946 Million cell

Title: US-09-720-384A-4
Perfect score: 1826
Sequence: 1 RPFHFNQTEPLVTHQTPP.....PKAMAKQVLCYLENGYLQA 343

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	502.5	27.5	610	2	US-08-879-561-10
2	495	27.1	624	2	US-08-879-561-3
3	495	27.1	624	2	US-08-879-561-9
4	97.5	5.3	655	4	US-09-347-833-4
5	96.5	5.3	620	4	US-09-442-100-11
6	96.5	5.3	673	4	US-09-196-387-8
7	96.5	5.3	949	4	US-09-196-387-10
8	96.5	5.3	1327	4	US-09-196-387-2
9	95	5.2	540	2	US-08-884-072-3
10	95	5.2	540	4	US-09-212-168-3
11	95	5.2	1548	1	US-08-463-092B-7
12	95	5.2	1548	2	US-08-460-907B-7
13	94.5	5.2	582	3	US-08-906-865-3
14	94	5.1	415	2	US-08-815-718-2
15	91.5	5.0	635	1	US-07-832-855-2
16	91.5	5.0	1426	4	US-09-136-574A-43
17	89.5	4.9	251	2	US-08-704-931-4
18	88	4.8	1271	1	US-08-095-734-2
19	88	4.8	1271	3	US-08-444-623-2
20	88	4.8	1271	3	US-08-471-869-2
21	88	4.8	1271	5	US-09-342-563-2
22	88	4.8	1271	4	PCT-US94-08267-2
23	88	4.8	1274	4	US-09-095-443-2
24	87.5	4.8	1487	2	US-08-760-489-2
25	87.5	4.8	1487	2	US-08-760-489-4
26	87.5	4.8	1487	4	US-09-185-373-2
27	87.5	4.8	1487	4	US-09-185-373-4

28	87	4.8	681	1	US-08-083-590A-18
29	87	4.8	681	2	US-08-346-128-37
30	87	4.8	681	3	US-08-532-384-18
31	87	4.8	1078	1	US-08-264-534-32
32	87	4.8	1078	1	US-08-083-590A-11
33	87	4.8	1078	1	US-08-465-500-32
34	87	4.8	1078	2	US-08-346-128-32
35	87	4.8	1078	3	US-08-532-384-11
36	87	4.8	1078	3	US-08-893-828-32
37	87	4.8	2556	1	US-08-083-590A-20
38	87	4.8	2556	3	US-08-532-384-20
39	86.5	4.7	340	4	US-09-360-779-2
40	86.5	4.7	490	4	US-08-472-028A-10
41	86.5	4.7	490	4	US-09-071-296-10
42	86.5	4.7	490	4	US-09-196-268-10
43	86.5	4.7	490	4	US-09-015-683-10
44	86.5	4.7	490	4	US-09-191-998-10
45	85	4.7	163	4	US-09-199-637A-335

ALIGNMENTS

RESULT 1
US-08-879-561-10
; Sequence 10 Application US/08879561
; Patent No. 5817482
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Hawkins, Phillip R.
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: DISEASE RELATED NUCLEOTIDE KINASES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/879,561
; FILING DATE: Herewith
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0325 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 610 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 705385
US-08-879-561-10

Qy 125 EAPHSPVKEKPMWGN-----IGKSTNLMHNCILGQSDRQKLLGQ-----KGCVVWITGLS 175
Db 2 EIFGSLCKVKLSNNAQWGMORATNTVYQAHVSNRRGQVGTGRGFCVWLTGLS 61
Qy 176 GSKSTLACALSRELHCRGHLTVYDGDNLRLHGLNRLSPKAEADRAENIRRVGEVAKLFA 235
Db 62 GAGKTVSMALAEYLCHGIPCTYLDGDNIRQLAKNKGFSFDEENVRVIAEVAKLEA 121
Qy 236 DAGVICIASLSPYRRDADR-----ALLPHSNFIEVFDLPKICEARDPKGLYKUAR 290
Db 122 DAGLVCTISFISPTQDRNARQIHGASLP---EFEVFDAPLHVCEQDVKGLYKKAR 178
Qy 291 TGKIKGTGIDDPPEPPINGEIVIKMKD---EECPSPKAMAKOVLCYLEE 337
Db 179 AGEIKGTGIDSEYEKPEAPELVLKTDSCDVND-----VOQVVELLQE 222

RESULT 4
US-09-347-833-4
; Sequence 4, Application US/09347833
; Patent No. 6294658
; GENERAL INFORMATION:
; APPLICANT: Famodu, Lavo O.
; APPLICANT: Odell, Joan T.
; TITLE OF INVENTION: Factors Involved in Gene Expression
; FILE REFERENCE: BB-1172
; CURRENT APPLICATION NUMBER: US/09/347,833
; CURRENT FILING DATE: 1999-07-02
; EARLIER APPLICATION NUMBER: 60/092,415
; EARLIER FILING DATE: July 10, 1998
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Microsoft Office 97
; SEQ ID NO. 4
; LENGTH: 655
; TYPE: PRT
; ORGANISM: Oryza sativa
US-09-347-833-4

Query Match 5.3%; Score 97.5; DB 4; Length 655;
Best Local Similarity 27.5%; Pred. No. 0.2;
Matches 44; Conservative 11; Mismatches 62; Indels 43; Gaps 8;
Qy 7 NOTEPLVHTQPPSPAP-----GPASQGRQGNLTLLSPTTLAVILVNPQRA-----PPV 57
Db 407 SOMRPV-----MPPSVAPRPMYPPGVPVGGQQLFYGGPPA-----FVNPQPGFGQHL 458
Qy 58 LPLGTPSDAPLALVI-----HGLTP-RSSHSSAGL-----ASDSGRE 95
Db 459 IPCMRESVGPINFMVPMVQGGQPPQRPAGRRAGTGGIQPPMPMGHQOMLPRGSRGGYRY 518
Qy 96 GEGRGARTHCRGIRGWRVRRRRNGAAP-----GEAPHSPV 131
Db 519 ASGRGMPDNASRGVGLVPSPEYEMGRMPLSDAGAPPQVPI 558

RESULT 5
US-09-442-100-11
; Sequence 11, Application US/09442100
; Patent No. 6359193
; GENERAL INFORMATION:
; APPLICANT: Xu, Tian
; APPLICANT: Tao, Wufan
; APPLICANT: Wang, Weiyl
; APPLICANT: Zhang, Sheng
; APPLICANT: Yu, Wan
; TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF LATS
; TITLE OF INVENTION: GENES AND METHODS BASED THEREON
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York

; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/442,100
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/411,111
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 6523-003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 620 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-09-442-100-11

Query Match 5.3%; Score 96.5; DB 4; Length 620;
Best Local Similarity 23.2%; Pred. No. 0.23;
Matches 61; Conservative 30; Mismatches 65; Indels 107; Gaps 16;
Qy 20 PSPAPCPASQG-----ORQGNLTLLSPTTLAVILVNPQRPVLPGLPSD-APLPALVI 73
Db 93 PSPAQRPTSGNSGQQQTYGYSAPMPS-----NTQTEFAPLPSTPTNNAPMPT--- 143
Qy 74 HGLTPRSSHSSAGLASDSGRREGRGARTHCRHGIRW-VRRRRNGAAPGEAPHSPVK 132
Db 144 --TTRSAHS-----WPLTSLRTASSAFCSATRGEC 173
Qy 133 EKPVM-----SNIGKSTNLMHNCILGQSDRQKLLGQGV-----WVITGLSGSKSTLA 183
Db 174 SDALLPLHPAVIGADT-----LFROSEMEQKLGETNDARRRESIWST-----AGRK--- 219
Qy 184 CALSRELHCRGHLTVYDGDNLRLHGLNRLSFKAEADRAENIRRV-----GEVAKLF- 234
Db 220 -----EGQYLRP-----LRTKOKPENYQTIKIGKAGFGEV-KLVQ 254
Qy 235 --ADAGVICIASLISP--YRRDR 253
Db 255 KKADGKVYAMKSLIKTEWFKKQ 277

RESULT 6
US-09-196-387-8
; Sequence 8, Application US/09196387
; Patent No. 6277613
; GENERAL INFORMATION:
; APPLICANT: de Lange, Titia
; APPLICANT: Smith, Susan
; TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRF1 AND METHODS
; TITLE OF INVENTION: OF USE THEREOF
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue, 4th Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:

;; TITLE OF INVENTION: OF USE THEREOF
;; NUMBER OF SEQUENCES: 12
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Klauber & Jackson
;; STREET: 411 Hackensack Avenue, 4th Floor
;; CITY: Hackensack
;; STATE: New Jersey
;; COUNTRY: USA
;; ZIP: 07601
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/196,387
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 09/095,225
;; FILING DATE: June 10, 1998
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Jackson Esq., David A.
;; REGISTRATION NUMBER: 26,742
;; REFERENCE/DOCKET NUMBER: 600-1-230 CIP1
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 201-487-5800
;; TELEFAX: 201-343-1684
;; TELEX: 133521
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1327 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; HYPOTHETICAL: NO
US-09-196-387-2

Query Match 5.3%; Score 96.5; DB 4; Length 1327;
Best Local Similarity 20.2%; Pred. No. 0.7;
Matches 78; Conservative 34; Mismatches 143; Indels 131; Gaps 15;

QY 4 HFNQTEPLVTHQPPSPAPGASQGRQGNLTLLSPPTPLAVILVNPQR-----53
Db 13 HHOQQLQ-APGASAPPPPPPLSLPLGAPGTTTASGLAPFASPRHGLALPEGDGS 71
QY 54 -----APPVLPGLTPSD-----APLPALVIHGLTPRS 80
Db 72 RDPPDRSPDPVCGTSCSTTSTICTVAAAPVVPVAVSTSSAAGVAPNPAGSGSNNSPSS 131
QY 81 SHSAGLASDSGRREG-----EGRGARTHCHRGIRWRRRRRNGAAGEAPHSPVKE 133
Db 132 SSGFTSSSSSPSGSLAESPEAGVSTAPLPGP-----AAGPGTG-----175
QY 134 KPVMSIGKSTNHLWNHNLIGQSDRKLLQKGCVVWITGLSGSKSTLACALSRELHCR 193
Db 176 VPAVSG---ALRELLEACRNGDVKRKL-----VDAANVNNAKDMAGKSSPLHFA 223
QY 194 GHLTYVLGDGNLRHLDLSFKAEDRAENIRVGEVAKLFADAGVICIASLISPPRROR 253
Db 224 A-----GFRKDVVEHLLQMGANVIARDGGGLIPL-----H 254
QY 254 DACRALPHSNFIEVDLPKICEARDPKG-----LYKLAITGRIKIGFTGI----D 301
Db 255 NACS--FGHAEVVS-----LLCQGADPNARDNWNYPLEHAAIKGRIDVCIVLLQHGA 306
QY 302 DPEYPPINGEIVKWKDEECSPKAM 327
Db 307 DPNIRTDGRSALDIAD---PSAKAV 329

RESULT 9

US-08-884-072-3
;; Sequence 3, Application US/08884072
;; Patent No. 5872234
;; GENERAL INFORMATION:
;; APPLICANT: Bandman, Olga
;; APPLICANT: Corley, Neil C.
;; APPLICANT: Guegler, Karl J.
;; TITLE OF INVENTION: HUMAN EXTRACELLULAR MATRIX PROTEINS
;; NUMBER OF SEQUENCES: 6
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Incyte Pharmaceuticals, Inc.
;; STREET: 3174 Porter Drive
;; CITY: Palo Alto
;; STATE: CA
;; COUNTRY: USA
;; ZIP: 94304
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette
;; COMPUTER: IBM Compatible
;; OPERATING SYSTEM: DOS
;; SOFTWARE: FastSEQ for Windows Version 2.0
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/884,072
;; FILING DATE: Herewith
;; CLASSIFICATION: 424
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER:
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Billings, Lucy J.
;; REGISTRATION NUMBER: 36,749
;; REFERENCE/DOCKET NUMBER: PF-0333 US
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 415-855-0555
;; TELEFAX: 415-845-4166
;; TELEX:
;; INFORMATION FOR SEQ ID NO: 3:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 540 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; IMMEDIATE SOURCE:
;; LIBRARY: BRAITUT13
;; CLONE: 1621777
US-08-884-072-3
Query Match 5.2%; Score 95; DB 2; Length 540;
Best Local Similarity 23.0%; Pred. No. 0.27;
Matches 52; Conservative 26; Mismatches 84; Indels 64; Gaps 13;
QY 1 RPFHF-----INQTEPLVTHQPPSPAPGASQGRQGNLTLLSPPTPLAVIL 48
Db 34 RPEHFQEVGYAAPSPPLSLRSLPM-----DHPDSSQHGGPPEGQSQ---VQPPPSQE---82
QY 49 VNPQAPVLPGLTPSD---APLP--ALVIHGLTPRSSH---SSAGLAS---DSGRREG 96
Db 83 ATPLQEKELLPAQLPAEKVEGPPPLQEAVALQKELPSLQHPNEQKGMFAPFGDQSHPEP 142
QY 97 EGRGARTHCH--RGICRWVRRRRRNGAAGEAPHSPVKEKPVMSNIGK-----STNLIWH 149
Db 143 ESNWAAHQCOODRSQGGW--GHRLDGFPFG-----RPSDNLNLICLPNRHVVYIG 191
QY 150 NCLIQSDRKLLQKGCVVWITGLSGSKSTLACALSRELHCRGH 195
Db 192 PWNLPQSSYSHLTRQ-----GETLNFLEIGYSRCCHCRSH 226
RESULT 10
US-09-212-168-3
;; Sequence 3, Application US/09212168
;; Patent No. 6303765
;; GENERAL INFORMATION:

APPLICANT: Bandman, Olga
APPLICANT: Corley, Neil C.
APPLICANT: Guegler, Karl J.
TITLE OF INVENTION: HUMAN EXTRACELLULAR MATRIX PROTEINS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/212,168
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/884,072
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0333 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 540 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRAITUT13
CLONE: 1621777
US-09-212-168-3

Query Match 5.2%; Score 95; DB 4; Length 540;
Best Local Similarity 23.0%; Pred. No. 0,27;
Matches 52; Conservative 26; Mismatches 84; Indels 64; Gaps 13;

QY 1 RPFHF-----INTEPLVTHQQPPSPAPGASOGQGRQGNWLLSPPTLAVIL 48
Db ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
34 RPEHFQEVGYAAPSPPLSRSLPM-----DHPDSSQHGPFEGSQ-----VQPPPSQE--- 82
QY 49 VNPQAPPVPLGLTPSD---APLP--ALVIHGLTPRSSH---SSAGLAS---DSGRREG 96
Db ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
83 ATPLQEQKLLPAQLPAKEVGPPLPQEAVPLOKELPSLOHNEKECMAPAPCDQSHPEP 142
QY 97 EGRGARTICH--RQIGRWRRRRNGAAGEAPHSPVKEKPVMSNICK-----STNLIWH 149
Db ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
143 ESWNAACHCQDDRSQGGW--GHRLDGFPFG-----RPSDNLNQLCLPNRQHVYVG 191
QY 150 NCLIGSQDRKLLGQKGVVWITGLSGSGKSTLACALSRLHCRGH 195
Db ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
192 PNWLQSSYSHLTRQ-----GETLNFLEIGYSRCCHRS 226

RESULT 11

US-08-463-092B-7
Sequence 7, Application US/08463092B
Patent No. 5766880
GENERAL INFORMATION:
APPLICANT: Cole, Susan P.C.
APPLICANT: Deeley, Roger G.
TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ENCODING
TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS

NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: PARTEQ RESEARCH & DEVELOPMENT INNOVATIONS
STREET: Queen's University at Kingston
CITY: Kingston
STATE: Ontario
COUNTRY: CANADA
ZIP: K7L 3N6
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/463,092B
FILING DATE: 05-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/966,923
FILING DATE: 27-OCT-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/029,340
FILING DATE: 8-MAR-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Steeg, Carol Miernicki
REGISTRATION NUMBER: 39,539
REFERENCE/DOCKET NUMBER: 01546
TELECOMMUNICATION INFORMATION:
TELEPHONE: (613) 545-2342
TELEFAX: (613) 545-6853
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1548 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: Internal
US-08-463-092B-7
Query Match 5.2%; Score 95; DB 1; Length 1548;
Best Local Similarity 21.3%; Pred. No. 1,2;
Matches 54; Conservative 41; Mismatches 101; Indels 58; Gaps 8;
QY 65 DAPLPALVHGLTPRSSHSSAGLASDSRREGGRCGARTHCHRGIRWVRRRRNGAAGP 124
Db ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
597 DASASSLAVHSTTVHMG-STQTVITDSGAGED-----EKEVEEG 637
QY 125 EAPHSPYKPKPVMSNIGKSTNLIWHNCLIGOSDRQKLLGQKGVVWITGLSGSGKSTLAC 184
Db ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
638 DREYIQLVSKELLRNV-----SLTIPKGLTWMVIGSTGSGKSTLLG 678
QY 185 AL-----SRELHCRGHLYVLDGDNLRHG-LNRDLSFEKAEDRAENIRRYGEVAKLFAD 236
Db ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
679 ALMGEYSVESGELWAERSATVYPOQAWIMNATLGNILFFDEERADLDVIRCCQLEAD 738
QY 237 AGVIC-----IASLISPYRRDR-DACRALLPHSNFIEVFDLPKICEARDPKGLY 286
Db ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
739 LAQFCGGLDTEIGEMGVNLGGQKARVSLARAV--YANRDVYLLDDPLSALDAHVGRIV 796
QY 287 KLARTGKIKGFTCI 300
Db ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
797 QDVILGRGRKTRV 810

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 2, 2002, 00:53:22 ; Search time 39 Seconds
(without alignments)
845.093 Million cell updates/sec

Title: US-09-720-384A-4
Perfect score: 1826
Sequence: 1 RPFHFNQTEPLVTHQTQPP.....PKAMAKOVLCLYLENGYLQA 343

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues
Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR-71.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match %	Score	Length	ID	Description
1	833	45.6	312	2 T08076	adenylisulfate ki
2	816	44.7	276	1 S47640	adenylisulfate ki
3	777.5	42.6	293	2 T06100	adenylisulfate ki
4	571	31.3	200	2 E96912	adenylisulfate ki
5	555	30.4	199	2 A69839	adenylisulfate ki
6	534.5	29.3	202	2 A83836	adenylisulfate ki
7	511.5	28.0	633	2 H95932	probable adenylis
8	510	27.9	213	2 AF0408	adenylisulfate ki
9	508.5	27.8	635	2 A87433	hypothetical prote
10	505	27.7	652	1 T24918	3'-phosphoadenosin
11	503.5	27.6	202	2 T50101	adenylisulfate ki
12	502.5	27.5	610	1 J64383	3'-phosphoadenosin
13	498.5	27.3	660	2 G2672	ATP sulfurylase, l
14	495	27.1	201	1 A00856	adenosine 5'-phosph
15	495	27.1	624	1 JW0087	3'-phosphoadenosin
16	494.5	27.1	615	2 F82062	adenylisulfate ki
17	492	26.9	644	2 A03471	sulfate adenylitr
18	490.5	26.9	641	1 Z22RNO	3'-phosphoadenosin
19	490.5	26.9	641	1 Z22RNO	adenylisulfate ki
20	481	26.3	201	1 B65056	adenylisulfate ki
21	481	26.3	201	2 D91079	adenosine 5'-phosp
22	481	26.3	201	2 E85924	adenosine 5'-phosp
23	481	26.3	208	2 A84073	adenylisulfate ki
24	480	26.3	197	2 C69877	adenylisulfate ki
25	478.5	26.2	620	1 I39755	3'-phosphoadenosin
26	474	26.0	202	1 S17244	adenylisulfate ki
27	457	25.0	206	2 H84978	adenylisulfate ki
28	448.5	24.6	196	2 H83472	adenosine 5'-phosp
29	402	22.0	614	1 B70772	probable 3'-phosph

30 398 21.8 546 1 C70393 probable 3'-phosph
31 387 21.2 177 1 S74917 adenylisulfate ki
32 369 20.2 214 1 S18729 adenylisulfate ki
33 351 19.2 192 2 B75594 adenylisulfate ki
34 325 17.8 186 2 G72590 probable adenylis
35 322.5 17.7 574 1 S55034 sulfate adenylitr
36 314.5 17.2 174 2 R75097 adenylisulfate 3-
37 307 16.8 573 1 A53651 sulfate adenylitr
38 305 16.7 155 2 H69285 ATP sulfurylase 3-
39 302 16.5 633 2 D83091 adenylisulfate 3-
40 179 9.8 170 2 G81286 probable adenylis
41 111.5 6.1 396 2 T36678 probable septum si
42 108 5.9 1016 2 D86308 translation initia
43 106 5.8 478 2 T10030 hypothetical prote
44 106 5.8 586 2 H86914 conserved hypotet
45 106 5.8 723 2 B38749 3-phosphatidylinos

ALIGNMENTS

RESULT 1
T08076
adenylisulfate kinase (EC 2.7.1.25) precursor ; Madagascar periwinkle
N:Alternate names: adenosine-5'-phosphosulfate-kinase
C:Species: Catharanthus roseus (Madagascar periwinkle)
C:Date: 21-May-1999 #sequence_revision 21-May-1999 #text_change 21-Jan-2000
C:Accession: T08076
R:Arz, H.E.; Gisselmann, G.; Schiffmann, S.; Schwenn, J.D.
Biochim. Biophys. Acta 1218, 447-452, 1994
A:Title: A cDNA for adenylsulfate kinase (APS)-kinase from Arabidopsis thaliana.
A:Reference number: S47640; MUID:94325358
A:Accession: T08076
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-312 <AR>
A:Cross-references: EMBL:AF044285; NID:g2832299; PIDN:AAC31145.1; PID:g2832300
C:Genetics:
A:Gene: aka
C:Function:
A:Description: phosphorylates 3'-OH group of adenosine-5'-phosphosulfate
C:Superfamily: adenylisulfate kinase; adenylisulfate kinase homology
C:Keywords: ATP; P-loop; phosphotransferase; purine nucleotide binding
F:135-298/Domain: adenylisulfate kinase homology <ASK>

Query Match	45.6%	Score 833;	DB 2;	Length 312;
Best Local Similarity	72.9%	Pred. No. 2.8e-56;		
Matches 156;	Conservative 24;	Mismatches 34;	Indels 0;	Gaps 0;
Qy	130	PVKEKPVMSNIGKSTNIIWHNCLIGQSDROKLLGQKGVVWITGLSGSGKSTLACALSRE	189	
Db	99	FGKKILQTTVGNSTNIIWHKCAVEKSERQEPQQRGCVITWITGLSGSGKSTLACALSRG	158	
Qy	190	LHCGRHHTYVLVDGDNLRHGLNRLDSFKAEDRAENIRRVGEVAKLFADAGVICIASLISPY	249	
Db	159	LHAKGLITYILDGDNVNRHGLNRLDSFKAEDRAENIRRVGEVAKLFADAGVICIASLISPY	218	
Qy	250	RRDRACRALLPHSNIEVFIDPLKICEARDPKGLYKARTGKIKGTGDDPYEPPIN	309	
Db	219	RKPPDRCRSLLEPGDFIEVFMDVPLKVCPEARDPKGLYKILARAGKIKGTGDDPYEPLK	278	
Qy	310	GEIVKMKDEECPSPKAMAKOVLCLYLENGYLQA	343	
Db	279	SEIVLHQKMGCDSPCDLADIVISYLENGYLKA	312	

RESULT 2
S47640
adenylisulfate kinase (EC 2.7.1.25) precursor - Arabidopsis thaliana
N:Alternate names: APS kinase; protein F26C24.11
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 02-Feb-2001
C:Accession: S47640; T02601; R84521; S38587

F:161-171/Domain: 3'-phosphoadenosine-5'-phosphosulfate binding #status predicted <P>
F:211-605/Domain: sulfate adenylyltransferase homology <SAT>
F:118/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 27.5%; Score 502.5; DB 1; Length 610;
Best Local Similarity 49.3%; Pred. No. 1.2e-30;
Matches 103; Conservative 34; Mismatches 51; Indels 21; Gaps 5;

QY 143 STNLIWHNCLIGQSDRQKLLGQ-----KGCVMWITGLSGSGKSTLACALSRELHCRHGLTY 198
DB 10 ATNVTQTQHVSRRAKQGVLGQGGFRGCTVWFTGLSGAGTTTISFALEYLVQSQIPT 69
QY 199 VLDGDNLRHGLNRLDSFKAEDRAENIRRVGEVAKLFADAGVICTIASLSPYRRDRDACRA 258
DB 70 SLDGDNVRHGLNKLNGFTQEDRENIIRLSEVAKLFDAGDGVICLTFSFPFKDRDLARS 129
QY 259 L-----LPHSNFIEVDLPLKICEARDPKGLYKLTARTGKFTGIDDDPYEPPINGEIV 313
DB 130 LHEQAGLEP---FFCFKFDVTDPLDVCQORDVKGGLYKARAGQIKGFTGIDQOYESPDAPETQ 186
QY 314 I---KMKDECPSPKAMAKQVLCYLEENG 339
DB 187 LVAGNKSIDEC-----VQEVWSLLQKNG 209

RESULT 13
G82672
ATP sulfurylase, large subunit XF1501 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Sep-2000
C:Accession: G82672
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A89515; MUID:20365717
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: G82672
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-660 <SIM>
A:Cross-references: GB:AE003980; GB:AE003849; NID:g9106531; PIDN:AAF84310.1; GSPDB:GN
A:Experimental source: strain 9a5c
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.;
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrier
as-Neto, E.; Docena, C.; El-borhy, H.; Facincan, A.P.; Ferreira, A.J.S.
submitted to Genbank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fr
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kutamae, E.E.; La
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Martino, C.L.; Marques, M.V.; Martins
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Mencia, C.F.M.; Miracca, E.C.; Miyaki, C.
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri,
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawa
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv
M.; Tsunako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.
A:Reference number: A59328
C:Genetics:
C:Contents: annotation
C:Gene: XF1501
C:Superfamily: nodulation protein nodQ; adenylylsulfate kinase homology; translation

Query Match 27.3%; Score 498.5; DB 2; Length 660;
Best Local Similarity 43.9%; Pred. No. 2.6e-30;
Matches 101; Conservative 42; Mismatches 76; Indels 11; Gaps 2;

QY 106 HRGIGRVRRRRRNGAAGPEAPSPKVPKPYMSNIGKSTNLIWHNCLIGQSDRQKLLGQ 165
DB 437 NRVLSGFILIDRHSNATVGAG-----TLDGSGVHRASNVHWPQIDIDHVARARIKQT 488
QY 166 GCVMWITGLSGSGKSTLACALSRELHCRHGLTYVLDGDNLRHGLNRLDSFKAEDRAENIR 225
DB 489 PKVILWFTGLSAGKASTIANIVDKRLHALGYHTFTILDGDNVRHGLNRLDSFTVEDVREIR 548
QY 226 RVGEVAKLFDAGVICTIASLSPYRRDRDACRALPHSNFTEVFDLPKICEARDPKGL 285

QY 298 TGIDDPYEPPIVKK-KD---EECPSPKAMAKQVLCYLEENG 341
DB 190 TGIDSAYPEPNAEIIIDAGDGQVQC-----VQKVLHDLESKGL 230

RESULT 11
T50101
adenylylsulfate kinase [imported] - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 21-Jul-2000
C:Accession: T50101
R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Brown, S.; Harris, D.
submitted to the EMBL Data Library, February 2000
A:Reference number: Z25037
A:Accession: T50101
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-202 <WOO>
A:Cross-references: EMBL:AL158056; PIDN:CA876273.1; GSPDB:GN00066; SPDB:SPAC1782.11
A:Experimental source: strain 972h(-); cosmid c1782
C:Genetics:
A:Gene: SPDB:SPAC1782.11
A:Map position: 1
A:Introns: 157/1
C:Superfamily: adenylylsulfate kinase; adenylylsulfate kinase homology

Query Match 27.6%; Score 503.5; DB 2; Length 202;
Best Local Similarity 51.7%; Pred. No. 2.5e-31;
Matches 105; Conservative 29; Mismatches 52; Indels 17; Gaps 4;

QY 143 STNLIWHNCLIGQSDRQKLLGQCVWITGLSGSGKSTLACALSRELHCRHGLTYVLDG 202
DB 2 ATNITHPGVSVKKEIRIKFVGHPCGTIWTGLSAGSKSTIACALEQVLLQRGVTTYRLDS 61
QY 203 DNLRHGLNRLDSFKAEDRAENIRRVGEVAKLFADAGVICTIASLSPYRRDRDACRAL--- 259
DB 62 DNVFGLNSDLGFSQDRNENIRIGHVAKLFADAGVAVTSFISPYRKORDQAREFHKK 121
QY 260 --LPHSNFIEVDLPLKICEARDPKGLYKLTARTGKIKGFTGIDDPYEPPIVKK-- 315
DB 122 DGLP-----FTEVVECPVAVQRDPKGLYKARAGEIKFTGISAPYEAPISPEIVVSSH 178
QY 316 -MKDECPSPKAMAKQVLCYLEE 337
DB 179 TQSIIEC-----VEKIVNYLLE 195

RESULT 12
JC4383
3'-phosphoadenosine-5'-phosphosulfate synthetase - spoonworm (Urechis caupo)
N:Alternate names: adenosine 5'-phosphosulfate kinase; PAPS
N:Alternates: adenylylsulfate kinase (EC 2.7.1.25); sulfate adenylyltransferase (EC 2.7.7.
C:Species: Urechis caupo
C:Date: 17-Jan-1996 #sequence_revision 19-Apr-1996 #text_change 19-Jan-2001
C:Accession: JC4383
R:Rosenthal, E.; Leustek, T.
Gene 165, 243-248, 1995
A:Title: A multifunctional Urechis caupo protein, PAPS synthetase, has both ATP sulfuryl
A:Reference number: JC4383; MUID:96096529
A:Accession: JC4383
A:Molecule type: mRNA
A:Residues: 1-610 <ROS>
A:Cross-references: GB:L39001; NID:g705384; PIDN:AAB00139.1; PID:g705385
C:Genetics:
A:Gene: paps
C:Function: <ASKF>
A:Description: as adenylylsulfate kinase catalyzes the phosphorylation of adenylylsulfat
C:Function: <SAFF>
A:Description: as sulfate adenylyltransferase catalyzes the reaction of sulfate and ATP
C:Superfamily: animal 3'-phosphoadenosine-5'-phosphosulfate synthetase; adenylylsulfate
C:Keywords: multifunctional enzyme; nucleotide binding; nucleotide binding; nucleotide binding; P-loop;
F:37-200/Domain: adenylylsulfate kinase homology <ASK>
F:44-51/Region: nucleotide-binding motif A (P-loop)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 2, 2002, 01:41:08 ; Search time 20 Seconds
(without alignments)
664.040 Million cell updates/sec

Title: US-09-720-384A-4
Perfect score: 1826
Sequence: 1 RPFHINQTEPLVTHQTPP.....PRAMAKVLCYLENGYIQA 343

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues
Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	833	45.6	312	1 KAPS_CATRO	Q49204 catharanthu
2	816	44.7	276	1 KAP1_ARATH	Q43295 arabidopsis
3	777.5	42.6	293	1 KAP2_ARATH	Q49196 arabidopsis
4	555	30.4	199	1 KAP2_BACSU	Q06735 bacillus su
5	534.5	29.3	202	1 CYC1_BACHD	Q9kct0 bacillus ha
6	505	27.7	211	1 KAPS_PENCH	Q12657 penicillium
7	502.5	27.5	633	1 NODQ_RHISB	Q07309 r nodg bifu
8	498.5	27.3	623	1 CYSN_XYLFA	Q9d78 x cyn/cysc
9	497.5	27.2	632	1 NODQ_RHTR	P52978 r nodg bifu
10	495	27.1	624	1 PPS1_HUMAN	Q43252 h bifunctio
11	495	27.1	624	1 PPS1_MOUSE	Q60967 m bifunctio
12	494.5	27.1	215	1 CYSN_VIBCH	Q9kp21 vibrio chol
13	490.5	26.9	641	1 NODQ_RHME	P13442 r nodg bifu
14	490	26.8	624	1 PPS1_CAVPO	Q54820 c bifunctio
15	486	26.6	206	1 KAPS_EMENI	Q92203 emericeila
16	482	26.4	614	1 PPS2_HUMAN	Q95340 h bifunctio
17	481.5	26.4	621	1 PPS2_MOUSE	Q88428 m bifunctio
18	481	26.3	200	1 CYSN_ECOLI	P23846 escherichia
19	481	26.3	208	1 CYC2_BACHD	Q9k7h6 bacillus su
20	480	26.3	197	1 CYC1_BACSU	Q34577 bacillus su
21	478.5	26.2	620	1 NODQ_AZOB	P28604 a nodg bifu
22	474	26.0	202	1 KAPS_YEAST	Q02196 saccharomyc
23	474	26.0	646	1 NODQ_RHIS3	P72339 r nodg bifu
24	457	25.0	206	1 CYSN_BUCAI	P57497 buchnera ap
25	448.5	24.6	196	1 CYC1_PSEAE	P57702 pseudomonas
26	402	22.0	614	1 CYSN_MYCTU	Q10600 m cyn/cysc
27	398	21.8	546	1 SATC_AQUAE	Q67174 a probable
28	387	21.2	177	1 CYSN_SYNY3	P72940 synechocyst
29	369	20.2	214	1 CYC2_PSEAE	P29811 pseudomonas
30	351	19.2	192	1 CYSN_DEIRA	P56861 deinococcus
31	325	17.8	186	1 CYSN_AERPE	Q9ycr6 aeropyrum p
32	322.5	17.7	574	1 MET3_EMENI	Q12555 emericeila
33	322	17.6	568	1 MET3_ASPTT	P56862 aspergillus

RESULT 1				
ID	KAPS_CATRO	STANDARD;	PRT;	312 AA.
AC	O49204;			
DT	30-MAY-2000 (Rel. 39, Created)			
DT	30-MAY-2000 (Rel. 39, Last annotation update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Adenylylsulfate kinase, chloroplast precursor (EC 2.7.1.25) (APS kinase) (Adenosine-5-phosphosulfate kinase) (ATP adenosine-5'-phosphosulfate 3'-phosphotransferase).			
DE	kinase) (Adenosine-5-phosphosulfate kinase) (ATP adenosine-5'-phosphosulfate 3'-phosphotransferase).			
GN	AKN.			
OS	Catharanthus roseus (Rosy periwinkle) (Madagascar periwinkle).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Gentianales; Apocynaceae; Rauvolfioideae; Vinaceae; Catharanthus.			
OC	Vinaceae; Catharanthus.			
OX	NCBI_TaxID=4058;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Schiffmann S., Schwenn J.-D.;			
RT	"Isolation of cDNA clones encoding adenosine-5'-phosphosulfate-kinase (EC2.7.1.25) from Catharanthus roseus and an isoform (akn2) from Arabidopsis.";			
RT	Arabidopsis.";			
RL	(In) Plant Gene Register PCR98-116.			
CC	-1- FUNCTION: CATALYZES THE SYNTHESIS OF ACTIVATED SULFATE.			
CC	-1- CATALYTIC ACTIVITY: ATP + adenylylsulfate = ADP + 3'-phosphoadenylylsulfate.			
CC	-1- PATHWAY: SECOND STEP IN THE SULFATE ACTIVATION PATHWAY. THIS REACTION OCCURS EARLY IN THE REDUCTIVE BRANCH OF THE CYSTEINE BIOSYNTHETIC PATHWAY.			
CC	-1- SUBCELLULAR LOCATION: Chloroplast (By similarity).			
CC	-1- SIMILARITY: BELONGS TO THE APS KINASE FAMILY.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; AF044285; AAC31145.1; --			
DR	Mendel; 28081; Catro;1772;28081.			
DR	InterPro; IPR002891; APS_kinase.			
DR	Pfam; PF01583; APS_kinase; 1.			
DR	ProDom; PD002350; APS_Kinase; 1.			
KW	Transferase; Kinase; Cysteine biosynthesis; ATP-binding; Phosphorylation; Transit peptide; Chloroplast.			
FT	TRANSIT 1 ?			
FT	CHAIN ?			
FT	NP_BIND 142 149			
FT	ACT_SITE 216 216			
FT	FORMS THE PHOSPHOSERINE INTERMEDIATE (BY SIMILARITY).			
SQ	SEQUENCE 312 AA; 33656 MW; 16BBD11FB4B1FE27 CRC64;			
Query Watch 45.6%; Score 833; DB 1; Length 312;				

P56858 pyrococcus
O29953 archaeoglob
Q12650 penicillium
O50274 p cyn/cysc
O9sh11 arabidopsis
R23726 bos taurus
Q15032 homo sapien
Q9fpq6 chlamydomon
P11675 pseudorabie
Q92777 homo sapien
O63537 rattus norv
Q92519 mycobacteri

ALIGNMENTS

Best Local Similarity 72.9%; Pred. No. 5.5e-55;
Matches 156; Conservative 24; Mismatches 34; Indels 0; Gaps 0;

QY 130 PVKEKPMNSIGKSTNLHNCILGSDROKLLGQKGVWITGLSGSKSTLACALSRE 189
DB 99 PGKTLQTTVGNSTNLHNCILGSDROKLLGQKGVWITGLSGSKSTLACALSRG 158
QY 190 LHCGRHTYVLDDGNLRHGLNRLSFKAEADRAENIRRVGEVAKLFADAGVICIASLISPY 249
DB 159 LHAKGLTYILDDGNVRHGLNSDLSFKAEADRAENIRRVGEVAKLFADAGVICIASLISPY 218
QY 250 RDRDACPALLPHSNTEVFDLPLKICEADOPKGLYKARTCKIKFTGIDDPYEPPEIN 309
DB 219 RRPDDKSLLPEDGTEVMDVPLKVCEDRDPKGLYKARAKIKFTGIDDPYEPPLK 278
QY 310 GRIVTKMKDEECPSPKAMQVLCYLENGYLQA 343
DB 279 SEIVLHQKLGKCDSPCDADIVISYLENGYLKA 312

RESULT 2

KAPL_ARATH ID KAPL_ARATH STANDARD; PRT; 276 AA.
AC Q43295;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Adenylsulfate kinase 1, chloroplast precursor (EC 2.7.1.25) (APS
kinase) (Adenosine-5'-phosphosulfate kinase) (ATP adenosine-5'-
phosphosulfate 3'-phosphotransferase)
DE AKN1 OR AT2G14750 OR F26C24.11 OR T6B13.1.
GN Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID:3702;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Leaf.
RX MEDLINE=943295359; PubMed=8049272;
RA Arz H.E., Gisselmann G., Schiffmann S., Schwenn J.-D.;
RT "A cDNA for adenyl sulphate (APS)-kinase from Arabidopsis
thaliana.";
RL Biochim. Biophys. Acta 1218:447-452(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94345022; PubMed=8066145;
RA Jain A., Leustek T.;
RT "A cDNA clone for 5'-adenylphosphosulfate kinase from Arabidopsis
thaliana.";
RL Plant Physiol. 105:771-772(1994).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=CV. LANDSBERG ERECTA;
RA Leustek T.;
RT "Analysis of ATP-sulphurylase and APS-kinase genomic DNA sequences
in Arabidopsis thaliana.";
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RX STRAIN=CV. COLUMBIA;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
Fuji C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H.L.,
Moffat K.S., Cronin L.A., Shen M., Pai G., Van Aken S., Umayam L.,
Tallon L.J., Gill J.E., Adams M.D., Carrera A.J., Creasy T.H.,
Goodman H.M., Somerville C.R., Copenhaver G.P., Preuss D.,
Nierman W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,
Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
thaliana.";
RL Nature 402:761-768(1999).

CC -!- FUNCTION: CATALYZES THE SYNTHESIS OF ACTIVATED SULFATE.
CC -!- CATALYTIC ACTIVITY: ATP + adenylsulfate = ADP + 3'-
phosphoadenylsulfate.
CC -!- PATHWAY: SECOND STEP IN THE SULFATE ACTIVATION PATHWAY. THIS
REACTION OCCURS EARLY IN THE REDUCTIVE BRANCH OF THE CYSTEINE
BIOSYNTHETIC PATHWAY.
CC -!- SUBCELLULAR LOCATION: Chloroplast.
CC -!- SIMILARITY: BELONGS TO THE APS KINASE FAMILY.
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X75782; CAA53426.1; -
DR EMBL: U05238; AAC50035.1; -
DR EMBL: U59759; AAC50034.1; -
DR EMBL: AC004705; AAC24182.1; -
DR InterPro: IPR002891; APS_kinase.
DR Pfam: PF01583; APS_kinase; 1.
DR ProDom: PD002350; APS_kinase; 1.
KW Transferase; Kinase; Cysteine biosynthesis; ATP-binding;
KW Phosphorylation; Transit peptide; Chloroplast.
FT TRANSIT 1 37 CHLOROPLAST (POTENTIAL).
FT CHAIN 38 276 ADENYLSULFATE KINASE 1.
FT NP_BIND 108 115 ATP (POTENTIAL).
FT ACT_SITE 182 182 FORMS THE PHOSPHOSERINE INTERMEDIATE (BY
SIMILARITY).
FT SEQUENCE 276 AA; 29787 MW; CB698643AA09D811 CRC64;
Query Match 44.7%; Score 816; DB 1; Length 276;
Best Local Similarity 64.4%; Pred. No. 8.7e-54;
Matches 159; Conservative 32; Mismatches 40; Indels 16; Gaps 4;
QY 100 GARTHCGRIGRWRRRRNGAAGPAP---HSPVKEKPMNSIGKSTNLHNCILGQS 156
DB 43 GSOTLSH-----NKGSIPEVKSINGTGQKGP-LSTVGNSTNIKWHCESEKV 91
QY 157 DRKLLGQKGVVITGLSGSKSTLACALSRELHCRGHLLTVLDGDLNRLHGLNRLSFK 216
DB 92 DRQRLDQKGVIVTGLSGSKSTLACALNQMLYQKGLCYLDGDLNRLHGLNRLSFK 151
QY 217 AEDRAENIRRVGEVAKLFADAGVICIASLISPYRRDRACRALLPHSNFIEVFDLPLKI 276
DB 152 AEDRAENIRRVGEVAKLFADAGVICIASLISPYRTDRACRSLLEPGDFVEVNDVPLSV 211
QY 277 CEARDPKGLYKLARTGKIKFTGIDDPYEPPEINGEIVIKMKDEECPSPKANAKQVLCYLE 336
DB 212 CEARDPKGLYKLARAGKIKFTGIDDPYEPPLNCE--ISLGREGGTSPIEMAEKVGYL 269
QY 337 ENGYLQA 343
DB 270 NKGYLQA 276
RESULT 3
KAP2_ARATH ID KAP2_ARATH STANDARD; PRT; 293 AA.
AC O49196;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Adenylsulfate kinase 2, chloroplast precursor (EC 2.7.1.25) (APS
kinase) (Adenosine-5'-phosphosulfate kinase) (ATP adenosine-5'-
phosphosulfate 3'-phosphotransferase).
DE AKN2 OR AT4G39940 OR T5J17.110.
GN Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.

RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Hohelsel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Krieger J.E., Kuramae E.E., Laigret F., Lambas M.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA de Oliveira B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silva A.P., Terenzi M.F., Truffi D., Tsai S.M., Tshako M.H.,
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Maidanis J., Setubal J.C.;
RT "The genome sequence of the plant pathogen *Xylella fastidiosa*.";
RL Nature 406:151-159(2000).
CC -!- FUNCTION: ATP SULFURYLASE MAY BE THE GTPASE, REGULATING ATP
CC SULFURYLASE ACTIVITY (BY SIMILARITY).
CC -!- FUNCTION: APS KINASE CATALYZES THE SYNTHESIS OF ACTIVATED SULFATE
CC (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: ATP + sulfate = diphosphate + adenylylsulfate.
CC -!- CATALYTIC ACTIVITY: ATP + adenylylsulfate = ADP + 3'-
CC phosphoadenylylsulfate.
CC -!- PATHWAY: FIRST AND SECOND STEPS IN THE REDUCTIVE BRANCH OF THE
CC CYSTEINE BIOSYNTHETIC PATHWAY.
CC THESE REACTIONS OCCURS EARLY IN THE REDUCTIVE BRANCH OF THE
CC -!- SUBUNIT: HETERODIMER COMPOSED OF CYSD, THE SMALLER SUBUNIT, AND
CC CYSN (BY SIMILARITY).
CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE GTP-BINDING
CC ELONGATION FACTOR FAMILY. CYSN/NODQ SUBFAMILY.
CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE APS KINASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE003980: AAP84310.1; ALT_INIT.
DR InterPro: IPR002891: APS_kinase.
DR InterPro: IPR000795: GTP_EFTU.
DR InterPro: IPR004161: GTP_EFTU_D2.
DR Pfam: PF01583: APS_kinase; 1.
DR Pfam: PF03144: GTP_EFTU; 1.
DR Pfam: PF03144: GTP_EFTU_D2; 1.
DR PRINTS: PR00315: ELONGATNFCF.
DR PRODOM: PD002350: APS_kinase; 1.
DR PROSITE: PS00301: EFATOR_GTP; 1.
KW Cysteine biosynthesis; Transferase; Nucleotidyltransferase;
KW GTP-binding; Kinase; ATP-binding; Multifunctional enzyme;
KW Complete proteome.
FT DOMAIN 1 450 SULFATE ADENYLYL TRANSFERASE.
FT DOMAIN 451 623 ADENYLYLSULFATE KINASE.
FT NP_BIND 23 30 GTP (BY SIMILARITY).
FT NP_BIND 102 106 GTP (BY SIMILARITY).
FT NP_BIND 157 160 GTP (BY SIMILARITY).
FT NP_BIND 459 466 ATP (POTENTIAL).
FT ACT_SITE 533 533 FORMS THE PHOSPHOSERINE INTERMEDIATE (BY
FT SIMILARITY).
FT SQ SEQUENCE 623 AA; 68277 MW; C20730A365B28E94 CRC64;
Query Match 27.3%; Score 498.5; DB 1; Length 623;
Best Local Similarity 43.9%; Pred. No. 8.5e-30;
Matches 101; Conservative 42; Mismatches 76; Indels 11; Gaps 2;
QY 106 HRCIGWVRRRRRNCAPGAPSPVKPVPVMSNIGKSTNINLWNCILGSDRQKLLGQK 165

Db 400 NRVLSGFILIDRHNSATVGAG-----TLDGSHRASNVHQPLDIDHVARARIKQT 451
QY 166 GCVVITITGSGSKSTFLACALSRELHCRGHTLVLDGDLNRHGLNRDLSFKAEDRAENIR 225
Db 452 PKVLFITGLSGAGSKANIVDKRLHALGYHTFILDGDNVRHGLNRDLSFTVEDRVENIR 511
QY 226 RVGEVAKLPADAGVTCIASLISPRRDRDACRALLPHSNFIEVFDLPDKICEARDPKGL 285
Db 512 RVAEVARLVMDAGLVVLSFISPFDRERQRLAREFAANEVEFVDVPLAVERDVKGL 571
QY 286 YKLARTCKIKGTGIDDPYEPPIGEIVIKMDECPSPKAMAKOVLCYL 335
Db 572 YAKARAGLTDTFGIDSPYEPHPPELHLR---ADQGTPEQLASQVLSFL 618
RESULT 9
NODQ_RHTR
ID NODQ_RHTR STANDARD: PRT; 632 AA.
AC P52978;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE NodQ bifunctional enzyme (Modulation protein Q) [Includes: Sulfate
DE adenylyltransferase subunit 1 (EC 2.7.7.4) (Sulfate adenylylate
DE transferase) (SAT) (ATP-sulfurylase large subunit); Adenylylsulfate
DE kinase (EC 2.7.1.25) (APS kinase) (ATP adenosine-5'-phosphosulfate 3'-
DE phosphotransferase)].
GN NODQ.
OS Rhizobium tropici.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=398;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CFN 299;
RX MEDLINE=96303535; PubMed=8755625;
RA Laeremans T., Caluwaerts I., Verreth C., Rogel M.A.,
RA Vanderleyden J., Martinez-Romero E.;
RT "Isolation and characterization of Rhizobium tropici Nod factor
RT sulfation genes.";
RL Mol. Plant Microbe Interact. 9:492-500(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CIAT899;
RX MEDLINE=97002746; PubMed=8850086;
RA Folch-Mallol J.L., Marroqui S., Sousa C., Manyani H., Lopez-Lara I.M.,
RA van der Drift K.M.G.M., Haverkamp J., Quinto C., Gil-Serrano A.,
RA Thomas-Oates J., Spalink H.P., Megias M.;
RT "Characterization of Rhizobium tropici CIAT899 nodulation factors: the
RT role of nodH and nodPQ genes in their sulfation.";
RL Mol. Plant Microbe Interact. 9:151-163(1996).
CC -!- FUNCTION: PROPOSED TO PROVIDE ACTIVATED SULFATE FOR TRANSFER TO
CC NOD FACTOR. ATP SULFURYLASE MAY BE THE GTPASE, REGULATING ATP
CC SULFURYLASE ACTIVITY (BY SIMILARITY).
CC -!- FUNCTION: APS KINASE CATALYZES THE SYNTHESIS OF ACTIVATED SULFATE
CC (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: ATP + sulfate = diphosphate + adenylylsulfate.
CC -!- CATALYTIC ACTIVITY: ATP + adenylylsulfate = ADP + 3'-
CC phosphoadenylylsulfate.
CC -!- SUBUNIT: SULFATE-ACTIVATING ENZYMES, NODP AND NODQ, MAY BE
CC PHYSICALLY ASSOCIATED.
CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE GTP-BINDING
CC ELONGATION FACTOR FAMILY. CYSN/NODQ SUBFAMILY.
CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE APS KINASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC

```
CC or send an email to license@lsb-sib.ch).
CC -----
DR EMBL; U47272; AAB08984.1; -.
DR EMBL; X87608; CAA60914.1; -.
DR InterPro; IPR002891; APS_Kinase.
DR InterPro; IPR000795; GTP_EFTU.
DR InterPro; IPR004161; GTP_EFTU_D2.
DR Pfam; PF01583; APS_kinase; 1.
DR Pfam; PF00009; GTP_EFTU; 1.
DR Pfam; PF03144; GTP_EFTU_D2; 1.
DR ProDom; PD002350; APS_Kinase; 1.
DR ProSite; PS00301; EFATOR.GTP; 1.
KW Nucleotide; Transferase; Nucleotidyltransferase; GTP-binding;
KW Kinase; ATP-binding; Multifunctional enzyme.
FT DOMAIN 1 457 SULFATE ADENYL TRANSFERASE.
FT NP_BIND 458 632 ADENYL SULFATE KINASE.
FT NP_BIND 31 38 GTP (BY SIMILARITY).
FT NP_BIND 109 113 GTP (BY SIMILARITY).
FT NP_BIND 164 167 GTP (BY SIMILARITY).
FT ACT_SITE 466 473 ATP (POTENTIAL).
FT ACT_SITE 540 540 FORMS THE PHOSPHOSERINE INTERMEDIATE (BY
FT SIMILARITY).
SQ SEQUENCE 632 AA; 69814 MW; 8FA3DF0CC7C604B6 CRC64;
Query Match 27.2%; Score 497.5; DB 1; Length 632;
Best Local Similarity 53.2%; Pred. No. 1e-29;
Matches 101; Conservative 23; Mismatches 63; Indels 3; Gaps 1;
QY 142 KSTNLIWHNCLIGSDRQKLLGQKGVWITGLSGSGKSTACALSRHLCRHLYVLD 201
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
435 RADNVHWQALDVKRAASLKNQKPAVLTGLSGSGKSTANALEALLHTCGKHYLLD 494
QY 202 GDLRLGLNRLDSFKAEDRAENIRRYGEVAKLFPADAGVICIASLISPYRRDRACRALLP 261
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
495 GDNVRHGLNRLDGTAVDENIRRYAEVAKLMADAGLVICISLSPFRDRRWARELWG 554
QY 262 HSNFIEVFIDLPLKICEARDPKGLYKLTGKIGTGTIDDPYPPPINGEIVTKMDEEC 321
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
555 EGEFIEIFVDTDECARDPKGLYKAFAGNIAFTGVSSPYEAPENPELHKTGMGP- 613
QY 322 PSPKAMAKOV 331
Db : : : : :
614 --PARLALQI 621
RESULT 10
PPSI_HUMAN STANDARD; PRT; 624 AA.
AC O43252; O43841; O75332;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Bifunctional 3'-phosphoadenosine 5'-phosphosulfate synthetase 1 (PAPS
DE synthetase 1) (PAPSS 1) (Sulfurylase kinase 1) (SK1) (SK 1)
DE [Includes: Sulfate adenylyltransferase (EC 2.7.7.4) (Sulfate adenylylate
DE transferase) (SAT) (Atp-sulfurylase); Adenylylsulfate kinase
DE (EC 2.7.1.25) (Adenylylsulfate 3'-phosphotransferase) (APS kinase)
DE (Adenosine-5'-phosphosulfate 3'-phosphotransferase) (3'-
DE phosphoadenosine-5'-phosphosulfate synthetase)].
CN PAPSS1 OR PAPSS OR ATPSK1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Tonsil;
RX MEDLINE=98236023; PubMed=9576487;
RA Girard J.-P., Baekkevold E.S., Amalric F.;
RT "Sulfation in high endothelial venules: cloning and expression of the
RT human PAPS synthetase.";
RL FASEB J. 12:603-612(1998).
FT [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal brain;
RX MEDLINE=98334672; PubMed=9668121;
RA Venkatachalam K.V., Akita H., Strott C.A.;
RT "Molecular cloning, expression, and characterization of human
RT bifunctional 3'-phosphoadenosine 5'-phosphosulfate synthase and its
RT functional domains.";
RL J. Biol. Chem. 273:19311-19320(1998).
FT [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=98312048; PubMed=9648242;
RA Yanagisawa K., Sakakibara Y., Suiko M., Takami Y., Nakayama T.,
RA Nakajima H., Takayanagi K., Natori Y., Liu M.-C.;
RT "cDNA cloning, expression, and characterization of the human
RT bifunctional ATP sulfurylase/adenosine 5'-phosphosulfate kinase
RT enzyme.";
RL Biosci. Biotechnol. Biochem. 62:1037-1040(1998).
FT [4]
RP MUTAGENESIS OF HIS-151; ASN-426; GLY-427 AND HIS-428.
RX MEDLINE=99115594; PubMed=9915785;
RA Venkatachalam K.V., Fuda H., Koonin E.V., Strott C.A.;
RT "Site-selected mutagenesis of a conserved nucleotide binding HXGH
RT motif located in the ATP sulfurylase domain of human bifunctional
RT 3'-phosphoadenosine 5'-phosphosulfate synthase.";
RL J. Biol. Chem. 274:2601-2604(1999).
CC -1- FUNCTION: BIFUNCTIONAL ENZYME WITH BOTH ATP SULFURYLASE AND APS
CC KINASE ACTIVITY, WHICH MEDIATES TWO STEPS IN THE SULFATE
CC ACTIVATION PATHWAY. THE FIRST STEP IS THE TRANSFER OF A SULFATE
CC GROUP TO ATP TO YIELD ADENOSINE 5'-PHOSPHOSULFATE (APS), AND THE
CC SECOND STEP IS THE TRANSFER OF A PHOSPHATE GROUP FROM ATP TO APS
CC YIELDING 3'-PHOSPHOADENYL SULFATE (PAPS); ACTIVATED SULFATE DONOR
CC USED BY SULFOTRANSFERASE). IN MAMMALS, PAPS IS THE SOLE SOURCE OF
CC SULFATE; APS APPEARS TO BE ONLY AN INTERMEDIATE IN THE SULFATE-
CC ACTIVATION PATHWAY. ALSO INVOLVED IN THE BIOSYNTHESIS OF SULFATED
CC L-SELECTIN LIGANDS IN ENDOTHELIAL CELLS.
CC -1- CATALYTIC ACTIVITY: ATP + sulfate = diphosphate + adenylylsulfate.
CC -1- CATALYTIC ACTIVITY: ATP + adenylylsulfate = ADP + 3'-
CC phosphoadenylylsulfate.
CC -1- ENZYME REGULATION: INHIBITED BY CHLORATE.
CC -1- PATHWAY: BELONGS TO THE SULFATE ASSIMILATION PATHWAY THAT LEADS TO
CC THE BIOSYNTHESIS OF CYSTEINE AND METHIONINE, AND TO THE SULFATION
CC OF PROTEINS, CARBOHYDRATES, LIPIDS, DRUGS AND XENOBIOTICS.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN TESTIS, PANCREAS, KIDNEY, THYMUS,
CC PROSTATE, OVARY, SMALL INTESTINE, COLON, LEUKOCYTES AND LIVER.
CC ALSO EXPRESSED IN HIGH ENDOTHELIAL VENULES (HEV) CELLS AND IN
CC CARTILAGE.
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE APS KINASE
CC FAMILY.
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE SULFATE
CC ADENYL TRANSFERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@lsb-sib.ch).
CC -----
DR EMBL; Y10387; CAA71413.1; -.
DR EMBL; U53447; AAC39894.1; -.
DR EMBL; AF033026; AAC28429.1; -.
DR MIM; 603262; -.
DR InterPro; IPR002891; APS_Kinase.
DR InterPro; IPR002650; ATP-sulfurylase.
DR Pfam; PF01583; APS_kinase; 1.
DR Pfam; PF01747; ATP-sulfurylase; 1.
DR ProDom; PD002350; APS_kinase; 1.
DR ProDom; PD002381; Atp-sulfurylase; 1.
DR ProDom; PD002381; Atp-sulfurylase; 1.
DR Transferase; Nucleotidyltransferase; Kinase; Multifunctional enzyme;
KW ATP-binding; Multigene family.
FT DOMAIN 1 ?220 ADENYL SULFATE KINASE.
```

```
FT DOMAIN ?221 624 SULFATE ADENYLTRANSFERASE.
FT NP_BIND 59 66 ATP (POTENTIAL).
FT ACT_SITE 133 133 FORMS THE PHOSPHOSERINE INTERMEDIATE (BY SIMILARITY).
FT SITE 521 525 PP-MOTIF (BY SIMILARITY).
FT MUTAGEN 425 425 H->A: LOSS OF ACTIVITY.
FT MUTAGEN 426 426 N->K: INCREASED ACTIVITY.
FT MUTAGEN 427 427 G->A: 30% DECREASE IN ACTIVITY.
FT MUTAGEN 428 428 H->A: LOSS OF ACTIVITY.
FT MUTAGEN 427 428 GH->AA: LOSS OF ACTIVITY.
FT CONFLICT 270 270 L -> F (IN REF. 2).
FT CONFLICT 456 456 MISSING (IN REF. 2 AND 3).
FT CONFLICT 587 587 S -> L (IN REF. 2).
SQ SEQUENCE 624 AA; 70847 MW; 6BC4F9648016CA31 CRC64;

Query Match 27.1%; Score 495; DB 1; Length 624;
Best Local Similarity 46.1%; Pred. No. 1.5e-29;
Matches 106; Conservative 38; Mismatches 60; Indels 26; Gaps 6;

QY 125 EAPHSPVKEKPMNSN-----IGKSTNLMWHNCLIGSDRKLQ-----KGCVVWITGLS 175
DB 2 EIPGSLCKVKLSNNAQWGMQRATNVTYQAHVSRNKRQGVGTGRGCTVWITGLS 61
QY 176 GSGKSTLACALSRELHCRHGLTYVLDGDNLRHGLNRDLSFKAEDRAENIRRVGEVAKLFA 235
DB 62 GAGKTTVSMALAEYLVCHGIPCYTLDGDNIRQGLNKLGFSPEDREENVRIRAEVAKLFA 121
QY 236 DAGVICTIASLSPYRRDRACR-----ALLPHSNFIEVFDLPKICEARDPKGLYKLLAR 290
DB 122 DAGLVCTISFISPYTDNRNARQIHGASLP---FFEVFDAPLHVCEQDVKGLYKLLAR 178
QY 291 TGKIKGFTGIDDPYEPPIGEIVIKMKD---EECPSPKAMAKQVLCYLEE 337
DB 179 AGEIKGFTGIDSEYKPEAPELVKLTDSVDVND-----VQOVVELLQ 222

RESULT 11
PPSL_MOUSE STANDARD; PRT; 624 AA.
AC Q60967;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Bifunctional 3'-phosphoadenosine 5'-phosphosulfate synthetase 1 (PAPS synthetase 1) (PAPSS 1) (Sulfurylase kinase 1) (SKI) (SK 1)
DE [includes: Sulfate adenylyltransferase (EC 2.7.7.4) (Sulfate adenylyltransferase) (SAR) (Atp-sulfurylase); Adenylylsulfate kinase (EC 2.7.1.25) (Adenylylsulfate 3'-phosphotransferase) (APS kinase) (Adenosine-5'-phosphosulfate 3'-phosphotransferase) (3'-phosphoadenosine-5'-phosphosulfate synthetase)].
DE GN PAPSS1 OR PAPSS OR ASAPK OR ATPSK1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RN SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=96094345; PubMed=7493984;
RA Li H., Deyrup A., Mensch J.R. Jr., Domowicz M., Konstantinidis A.K., Schwartz N.B.;
RA "The isolation and characterization of cDNA encoding the mouse bifunctional ATP sulfurylase-adenosine 5'-phosphosulfate kinase.";
RL J. Biol. Chem. 270:29453-29459(1995).
CC -!- FUNCTION: BIFUNCTIONAL ENZYME WITH BOTH ATP SULFURYLASE AND APS KINASE ACTIVITY, WHICH MEDIATES TWO STEPS IN THE SULFATE ACTIVATION PATHWAY. THE FIRST STEP IS THE TRANSFER OF A SULFATE GROUP TO ATP TO YIELD ADENOSINE 5'-PHOSPHOSULFATE (APS), AND THE SECOND STEP IS THE TRANSFER OF A PHOSPHATE GROUP FROM ATP TO APS YIELDING 3'-PHOSPHOADENYLYLSULFATE (PAPS). ACTIVATED SULFATE DONOR USED BY SULFOTRANSFERASE). IN MAMMALS, PAPS IS THE SOLE SOURCE OF SULFATE; APS APPEARS TO BE ONLY AN INTERMEDIATE IN THE SULFATE-ACTIVATION PATHWAY. ALSO INVOLVED IN THE BIOSYNTHESIS OF SULFATED
```

```
CC L-SELECTIN LIGANDS IN ENDOTHELIAL CELLS.
CC -!- CATALYTIC ACTIVITY: ATP + sulfate = diphosphate + adenylylsulfate.
CC -!- CATALYTIC ACTIVITY: ATP + adenylylsulfate = ADP + 3'-phosphoadenylylsulfate.
CC -!- PATHWAY: BELONGS TO THE SULFATE ASSIMILATION PATHWAY THAT LEADS TO THE BIOSYNTHESIS OF CYSTEINE AND METHIONINE, AND TO THE SULFATION OF PROTEINS, CARBOHYDRATES, LIPIDS, DRUGS AND XENOBIOTICS.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN THE NEONATAL BRAIN AND IN CARTILAGE.
CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE APS KINASE FAMILY.
CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE SULFATE ADENYLTRANSFERASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).
CC -----
CC EMBL; U34883; AAC52328.1; .
CC MGD; MG1:1330587; Papss1.
CC InterPro; IPR002891; APS_kinase.
CC InterPro; IPR002650; Atp-sulfurylase.
CC Pfam; PF01583; APS_kinase; 1.
CC Pfam; PF01747; Atp-sulfurylase; 1.
CC ProDom; PD002350; APS_kinase; 1.
CC ProDom; PD002381; Atp-sulfurylase; 1.
CC Transferrase; Nucleotidyltransferase; Kinase; Multifunctional enzyme;
KW ATP-binding; Multigene family.
FT DOMAIN 1 2220 ADENYLTRANSFERASE.
FT NP_BIND 59 66 ATP (POTENTIAL).
FT ACT_SITE 133 133 FORMS THE PHOSPHOSERINE INTERMEDIATE (BY SIMILARITY).
FT SITE 521 525 PP-MOTIF.
SQ SEQUENCE 624 AA; 70794 MW; B487EFAF9B78BE3E CRC64;

Query Match 27.1%; Score 495; DB 1; Length 624;
Best Local Similarity 46.1%; Pred. No. 1.5e-29;
Matches 106; Conservative 38; Mismatches 60; Indels 26; Gaps 6;

QY 125 EAPHSPVKEKPMNSN-----IGKSTNLMWHNCLIGSDRKLQ-----KGCVVWITGLS 175
DB 2 EIPGSLCKVKLSNNAQWGMQRATNVTYQAHVSRNKRQGVGTGRGCTVWITGLS 61
QY 176 GSGKSTLACALSRELHCRHGLTYVLDGDNLRHGLNRDLSFKAEDRAENIRRVGEVAKLFA 235
DB 62 GAGKTTVSMALAEYLVCHGIPCYTLDGDNIRQGLNKLGFSPEDREENVRIRAEVAKLFA 121
QY 236 DAGVICTIASLSPYRRDRACR-----ALLPHSNFIEVFDLPKICEARDPKGLYKLLAR 290
DB 122 DAGLVCTISFISPYTDNRNARQIHGASLP---FFEVFDAPLHVCEQDVKGLYKLLAR 178
QY 291 TGKIKGFTGIDDPYEPPIGEIVIKMKD---EECPSPKAMAKQVLCYLEE 337
DB 179 AGEIKGFTGIDSEYKPEAPELVKLTDSVDVND-----VQOVVELLQ 222

RESULT 12
CYSC_VIBCH STANDARD; PRT; 215 AA.
AC Q9KPF21;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Adenylylsulfate kinase (EC 2.7.1.25) (APS kinase) (Adenosine-5'-phosphosulfate kinase) (ATP adenosine-5'-phosphosulfate 3'-phosphotransferase).
DE GN CYSC OR VC2558.
OS Vibrio cholerae.
```

OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrrio.
RX NCBI_TaxID=666;
[1]
PC SEQUENCE FROM N.A.
RX STRAIN=EL TOR N16961 / SROTYPE O1;
MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Emolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen *Vibrio*
RT cholerae.";
RL Nature 406:477-483(2000).
CC -1- FUNCTION: CATALYZES THE SYNTHESIS OF ACTIVATED SULFATE.
CC -1- CATALYTIC ACTIVITY: ATP + adenylylsulfate = ADP + 3'-
CC phosphoadenylylsulfate.
CC -1- PATHWAY: SECOND STEP IN THE SULFATE ACTIVATION PATHWAY. THIS
CC REACTION OCCURS EARLY IN THE REDUCTIVE BRANCH OF THE CYSTEINE
CC BIOSYNTHETIC PATHWAY.
CC -1- SIMILARITY: BELONGS TO THE APS KINASE FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF004324; AAF569.1; -
DR TIGR; VC2558; -
DR InterPro: IPR002891; APS_Kinase.
DR Pfam; PF01583; APS_Kinase; 1.
DR ProDom; PD002350; APS_Kinase; 1.
KW Transferase; Kinase; Cysteine biosynthesis; ATP-binding;
KW Phosphorylation; Complete proteome.
FT NP_BIND 46 53 ATP (BY SIMILARITY).
FT ACT_SITE 120 120 FORMS THE PHOSPHORINE INTERMEDIATE (BY
FT SIMILARITY).
SQ SEQUENCE 215 AA; 23906 MW; 83B9EE2F295CDDF6 CRC64;
Query Match 27.18; Score 494.5; DB 1; Length 215;
Best Local Similarity 48.38; Pred. No. 4.8e-30;
Matches 97; Conservative 36; Mismatches 65; Indels 3; Gaps 1;
QY 142 KSTNLHNCILIGQSDROKLLGQKGVVWITGLSGSGKSTLACALSRELHCRGLHYVLD 201
DB 15 KPNVWVRHRAVDKQKQATLQKQRPVWFTGLSGAGKSTVAGALENLALGYHTYLLD 74
QY 202 GDNLRHGLNRDLSFKAEDRAENIRRGVEAKLPADAGVICIASLISPYRRDRACRALLP 261
DB 75 GDNVRRHGLSGLDFSEQDRRENRIRIGELAKLSMDAGLIVLTAFTISPHRAERQMRDILLP 134
QY 262 HSNFIEVFDLPKICEARDPKGLYKLGRTGKIGFTGDDPYEPNGEIVLKMKDEEC 321
DB 135 NGELEVVYNTSLDVEARDPKGLYKKARAGEIRQFTGDSAYEAPLNPDLDLPAKEG-- 192
QY 322 PSKMAKQVLCVLENGYLQ 342
DB 193 -SVDELVAOCLQALAEHRIQ 212
RESULT 13
NODQ_RHIME
ID NODQ_RHIME STANDARD; PRT; 641 AA.
AC P13442;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE NodQ bifunctional enzyme (Nodulation protein Q) [Includes: Sulfate

DE DE adenylyltransferase subunit 1 (EC 2.7.7.4) (Sulfate adenylate
DE DE transferase) (SAT) (ATP-sulfurylase large subunit); Adenylylsulfate
DE DE kinase (EC 2.7.1.25) (APS kinase) (ATP adenosine-5'-phosphosulfate 3'-
DE DE phosphotransferase).
GN NODQ OR RA0469 OR SMA0857.
OS Rhizobium melliloti (Sinorhizobium melliloti).
OG Plasmid pSymA (megaplasmid 1).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RCR2011 / SU47;
RX MEDLINE=89313304; PubMed=2546009;
RA Cervantes E., Sharma S.B., Maillet F., Vasse J., Truchet G.,
RA Rosenberg C.;
RT "The Rhizobium melliloti host range nodQ gene encodes a protein which
RT shares homology with translation elongation and initiation factors.";
RL Mol. Microbiol. 3:745-755(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92288403; PubMed=2520820;
RA Schwedock J., Long S.R.;
RT "Nucleotide sequence and protein products of two new nodulation genes
RT of Rhizobium melliloti, nodP and nodQ.";
RL Mol. Plant Microbe Interact. 2:181-194(1989).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396509; PubMed=11481432;
RA Barnett M.J., Fisher R.F., Jones T., Komp C., Abola A.P.,
RA Barlett-Hubler F., Bowser L., Capela D., Galibert F., Gouzy J.,
RA Gurjal M., Hong A., Huizar L., Hyman R.W., Kahn D., Kahn M.L.,
RA Kelman S., Keating D.H., Palm C., Peck M.C., Surzycki R., Wells D.H.,
RA Yeh K.-C., Davis R.W., Federspiel N.A., Long S.R.;
RT "Nucleotide sequence and predicted functions of the entire
RT Sinorhizobium melliloti pSymA megaplasmid.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9883-9888(2001).
CC -1- FUNCTION: PROPOSED TO PROVIDE ACTIVATED SULFATE FOR TRANSFER TO
CC NOD FACTOR. ATP SULFURYLASE MAY BE THE GTPASE, REGULATING ATP
CC SULFURYLASE ACTIVITY (BY SIMILARITY).
CC -1- FUNCTION: APS KINASE CATALYZES THE SYNTHESIS OF ACTIVATED SULFATE
CC (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ATP + sulfate = diphosphate + adenylylsulfate.
CC -1- CATALYTIC ACTIVITY: ATP + adenylylsulfate = ADP + 3'-
CC phosphoadenylylsulfate.
CC -1- SUBUNIT: SULFATE-ACTIVATING ENZYMES, NODP AND NODQ, MAY BE
CC PHYSICALLY ASSOCIATED (POTENTIAL).
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE GTP-BINDING
CC ELONGATION FACTOR FAMILY. CYSN/NODQ SUBFAMILY.
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE APS KINASE
CC FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X14809; CAA32914.1; -
DR EMBL; M68859; AAA26343.1; -
DR EMBL; AE007237; AAK65127.1; -
DR PIR; S14899; ZZZRNO.
DR HSSP; P02990; 1ETU.
DR InterPro: IPR002891; APS_Kinase.
DR InterPro: IPR00795; GTP_EFTU.
DR InterPro: IPR004161; GTP_EFTU_D2.
DR Pfam; PF01583; APS_Kinase; 1.
DR Pfam; PF00009; GTP_EFTU; 1.
DR Pfam; PF03144; GTP_EFTU_D2; 1.
DR ProDom; PD002350; APS_Kinase; 1.

```
DR PROSITE; PS00301; EFACOR_GTP; 1.
KW plasmid; Nodulation; Transferase; Nucleotidyltransferase; GTP-binding;
KW Kinase; ATP-binding; Multifunctional enzyme; Complete proteome.
FT DOMAIN 1 458 SULFATE ADENYL TRANSFERASE.
FT DOMAIN 459 641 ADENYL SULFATE KINASE.
FT NP_BIND 31 38 GTP (BY SIMILARITY).
FT NP_BIND 110 114 GTP (BY SIMILARITY).
FT NP_BIND 165 168 GTP (BY SIMILARITY).
FT NP_BIND 467 474 ATP (POTENTIAL).
FT ACT_SITE 524 524 FORMS THE PHOSPHOSERINE INTERMEDIATE (BY
FT SIMILARITY).
SQ SEQUENCE 641 AA; 70614 MW; 1E1261F04ED33A93 CRC64;

Query Match 26.9%; Score 490.5; DB 1; Length 641;
Best Local Similarity 52.4%; Pred. No. 3.5e-29;
Matches 98; Conservative 25; Mismatches 63; Indels 1; Gaps 1;

QY 142 KSTNMLHNCILIGOSDQKLLGQGVVMTGLSGSKSTLACALSRELHCRGHLTYVLD 201
Db 436 RADNVHVALEVNKSARSAMKNQPAVLWFTGLSGSKSTIANELDRILHAQKHTYLLD 495

QY 202 GDNLRHGLNRLDSFKAEDRAENIRRVGEVAKLFADAGVICIASLISPYRRDRDCAALLP 261
Db 496 GDNVVRHGLNRLDGETEDRVENIRRVAEVAKLMADAGLIVLVSFSPFDRERMARELME 555

QY 262 HSNFIEVFDLPLKICEARDPKGLYKLGKIGFTGIDDPYEPPIGEIIVKMKDEEC 321
Db 556 EGFEIEIVDPLDECARDPKGLYEKALAGKIANFTGVSSCYEAPENPELHIRTVGHQ- 614

QY 322 PSPKAMA 328
Db 615 PNDLALA 621

RESULT 14
PPSL_CAVPO
ID PPSL_CAVPO STANDARD; PRT; 624 AA.
IC O54820:
DT 30-MAY-2000 (Rel. 39, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Bifunctional 3'-phosphoadenosine 5'-phosphosulfate synthetase 1 (PAPS
DE [includes: Sulfate adenylyltransferase (EC 2.7.7.4) (Sulfate adenylyl
DE transferase) (SAT) (ATP-sulfurylase); Adenylylsulfate kinase
DE (EC 2.7.1.25) (Adenylylsulfate 3'-phosphotransferase) (APS kinase)
DE (Adenosine-5'-phosphosulfate 3'-phosphotransferase) (3'-
DE phosphoadenosine-5'-phosphosulfate synthetase)].
DE PAPS1 OR PAPS OR ATPSK1.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NH 2;
RA Venkatchalam K.V., Akita H., Strott C.A.;
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: BIFUNCTIONAL ENZYME WITH BOTH ATP SULFURYLASE AND APS
CC KINASE ACTIVITY, WHICH MEDIATES TWO STEPS IN THE SULFATE
CC ACTIVATION PATHWAY. THE FIRST STEP IS THE TRANSFER OF A SULFATE
CC GROUP TO ATP TO YIELD ADENOSINE 5'-PHOSPHOSULFATE (APS), AND THE
CC SECOND STEP IS THE TRANSFER OF A PHOSPHATE GROUP FROM ATP TO APS
CC YIELDING 3'-PHOSPHOADENYL SULFATE (PAPS; ACTIVATED SULFATE DONOR
CC USED BY SULFOTRANSFERASE). IN MAMMALS, PAPS IS THE SOLE SOURCE OF
CC SULFATE; APS APPEARS TO BE ONLY AN INTERMEDIATE IN THE SULFATE-
CC ACTIVATION PATHWAY. ALSO INVOLVED IN THE BIOSYNTHESIS OF SULFATED
CC L-SELECTIN LIGANDS IN ENDOTHELIAL CELLS (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: ATP + sulfate = diphosphate + adenylylsulfate.
CC -!- CATALYTIC ACTIVITY: ATP + sulfate = diphosphate + adenylylsulfate.
CC -!- PATHWAY: BELONGS TO THE SULFATE METHYLATION PATHWAY THAT LEADS TO
CC THE BIOSYNTHESIS OF CYSTEINE AND METHIONINE, AND TO THE SULFATION
CC
```

```
CC OF PROTEINS, CARBOHYDRATES, LIPIDS, DRUGS AND XENOBIOTICS.
CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE APS KINASE
CC FAMILY.
CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE SULFATE
CC ADENYL TRANSFERASE FAMILY.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF004875; AAC02266.1; -.
CC InterPro; IPR002891; APS_Kinase.
CC InterPro; IPR002650; ATP-sulfurylase.
CC Pfam; PF01583; APS_Kinase; 1.
CC Pfam; PF01747; ATP-sulfurylase; 1.
CC ProDom; PD002350; APS_Kinase; 1.
CC ProDom; PD002381; ATP-sulfurylase; 1.
CC Transferase; Nucleotidyltransferase; Kinase; Multifunctional enzyme;
CC ATP-binding; Multigene family. ADENYL SULFATE KINASE.
FT DOMAIN 1 7220 SULFATE ADENYL TRANSFERASE.
FT NP_BIND 59 66 ATP (POTENTIAL).
FT ACT_SITE 133 133 FORMS THE PHOSPHOSERINE INTERMEDIATE (BY
FT SIMILARITY).
FT SITE 521 525 PP-MOTIF (BY SIMILARITY).
SQ SEQUENCE 624 AA; 70395 MW; BF7461B4D07F2131 CRC64;

Query Match 26.8%; Score 490; DB 1; Length 624;
Best Local Similarity 52.5%; Pred. No. 3.6e-29;
Matches 96; Conservative 30; Mismatches 45; Indels 12; Gaps 3;

QY 142 KSTNMLHNCILIGOSDQKLLGQ----KGCVVMTGLSGSKSTLACALSRELHCRGHLT 137
Db 24 RATNVTYQAHVHSRNKRGQVVGTRGFRGCTVMTGLSGAGKTTVSMALGHLVCHGIPC 83

QY 198 YVLDGDNLRHGLNRLDSFKAEDRAENIRRVGEVAKLFADAGVICIASLISPYRRDRDCA 257
Db 84 YTLGDNIRGKLNKNGSPEDRENVRRVIAEVAKLFDAGLVLCITFSIPTQDRNNAR 143

QY 258 -----ALLPHSFNIEVFDLPLKICEARDPKGLYKLGKIGFTGIDDPYEPPIGEI 312
Db 144 QIHEGASLP---FFEFVVDAPLHVCEORDVKGLYKARAGEIKGTGIDSEYKPEAPEL 200

QY 313 VIK 315
Db 201 VLK 203

RESULT 15
KAPS_EMENI
ID KAPS_EMENI STANDARD; PRT; 206 AA.
IC Q92203:
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE Adenylylsulfate kinase (EC 2.7.1.25) (APS kinase) (Adenosine-
DE 5'-phosphosulfate kinase) (ATP adenosine-5'-phosphosulfate 3'-
DE phosphotransferase).
DE GN SD.
OS Emericella nidulans (Aspergillus nidulans).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiiales; Trichocomaceae; Emericella.
OX NCBI_TaxID=5072;
RN [1]
RP SEQUENCE FROM N.A.
RA Clarke D.L., Newbert R.W., Turner G.;
RT "Cloning and characterisation of the APS kinase gene from
RT Aspergillus nidulans.";
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
RL
```

```
CC -!- FUNCTION: CATALYZES THE SYNTHESIS OF ACTIVATED SULFATE.
CC -!- CATALYTIC ACTIVITY: ATP + adenylylsulfate = ADP + 3'-
CC phosphoadenylylsulfate.
CC -!- PATHWAY: SECOND STEP IN THE SULFATE ASSIMILATION PATHWAY THAT
CC LEADS TO THE BIOSYNTHESIS OF METHIONINE AND CYSTEINE.
CC -!- SIMILARITY: BELONGS TO THE APS KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Y08866; CA470089.1; .
CC InterPro: IPR002891; APS_Kinase.
CC Pfam: PF01583; APS_Kinase; 1.
CC ProDom: PD002350; APS_Kinase; 1.
CC Transferrase; Kinase; Cysteine biosynthesis; ATP-binding;
CC phosphorylation.
CC NP_BIND 31 38 ATP (POTENTIAL).
CC ACT_SITE 105 105 FORMS THE PHOSPHOSERINE INTERMEDIATE (BY
CC SIMILARITY).
CC SEQUENCE 206 AA; 23028 MW; 6AADD483E2BCA1CD CRC64;

Query Match 26.6%; Score 486; DB 1; Length 206;
Best Local Similarity 50.7%; Pred. No. 2e-29;
Matches 106; Conservative 30; Mismatches 61; Indels 12; Gaps 5;

QY 143 STNLTWHNCLIGSGDROKLLGKGVVWITGLSGSGKSTLACALSRELHCRGHLTYVLDS 202
DB 2 ATNIT-HHAGLTNRNQLRKQKGLTIWLTGLSASGKSTIAVELEHQLLQRLHAYRLDG 60

QY 203 DNLRHGLNRDLSEKAEDRAENIRRVGEVAKLPADAGVICIASLSPYRRDRDACRAL--L 260
DB 61 DNVRFGLNKDLGFSADRNENIRRIAEVAKLFAFADSSIAITSFISPFADRDRTARKLHEV 120

QY 261 PHSN-----FIEVFIDLPLKICEARDPKGLYKLTARTKIGFTCIDDPYEPPIGEIIV 314
DB 121 PTPNDSTGLPFVEVVDVPIEVAEKRPDKGLYKKAREGIIKEFTGISSPYEAPENPEVHV 180

QY 315 KMRDECPSPKAMAKQVLCYLENGYLQA 343
DB 181 KNVD--LPIQEA--KQIIDIYLDKLLDA 206
```

Search completed: November 2, 2002, 01:52:18
Job time : 22 secs


```
Db 61 SYVLDGONLRHGLNKDGFKAEDRVENIRRVGEVAKLFADAGLICIASLISPYRKDRDAC 120
QY 257 RALLPHSNFTEVFDLPLKICARPDKGLYKARTGKIKGFTGIDDPYPPPINGEIVTKM 316
Db 121 REMTQNSFTEVFNMSLQCLCARPKGLYKLARAGKIKGFTGIDDPYSPINCLIELKE 180
QY 317 KDEBCPSPKAMAKQVLCYLEENGYLQ 342
Db 181 REGECPSPVAMAEVSYLEDKGFLQ 206

RESULT 2
Q9SE92
ID Q9SE92 PRELIMINARY; PRT; 288 AA.
AC Q9SE92;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE ADENOSINE-5'-PHOSPHOSULFATE KINASE (EC 2.7.1.25) (FRAGMENT).
GN AKI.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoidae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. DEKALB XL72;
RX MEDLINE=98403884; PubMed=9734815;
RA Kotani H., Nakamura Y., Sato S., Asamizu E., Kaneko T., Miyajima N.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. VI.
RT Sequence features of the regions of 1,367,185 bp covered by 19
RT physically assigned P1 and TAC clones.";
RL DNA Res. 5:203-216(1998).
DR EMBL; AB013390; BAB08460.1; -.
DR InterPro; IPR002891; APS_kinase.
DR Pfam; PF01583; APS_kinase; 1.
DR ProDom; PD002350; APS_kinase; 1.
KW Kinase; Transferase.
FT NON_TER 1
KW KINASE
SQ SEQUENCE 288 AA; 31612 MW; 291AE77AD184FF07 CRC64;

Query Match 44.5%; Score 812.5; DB 10; Length 288;
Best Local Similarity 53.9%; Pred. No. 4.9e-59;
Matches 172; Conservative 34; Mismatches 62; Indels 51; Gaps 7;

QY 40 TPPTLAVILVN-----PORAPPVPLGLTPSDAPLPALVIHGLTPRSSHSSAGLASDSGR 94
Db 2 VPVLLLVAVASRRAPPEPRVKSRSVASAPPPPE-----TP-----SSGN 43

QY 95 EGEGRGARTCHRGIRWYRRNRG-----AARGEAPHSPVKEK--PVMSNIGK 142
Db 44 QNEG-----RRKURGTPLVECTGDRSVEEQPEHAGVSGKAWKMSSTVPK 89

QY 143 STNLMHNCILIGSDRQKLLGQKGVVWITGLSGSGKSTLACALSRELHCRGHLTYVLDG 202
Db 90 SSNIFWHDPCVGTDRQNVLKQKGVVWITGLSGSGKSTLACTLGLRELHTRGKLAYVLDG 149

QY 203 DNLRHGLNRDLSEKAEDRAENIRRVGEVAKLFADAGVICIASLISPYRRDRDACLPH 262
Db 150 DNLRHGLNRDLSEKAEDRAENIRRV--VAKLFADAGVICIASLISPHRRDRESCALLSD 207

QY 263 SNFTEVFDLPLKICARPDKGLYKARTGKIKGFTGIDDPYPPPINGEIVKMKDECP 322
Db 208 SSFTEVFNLSLELCEARDPKGLYKLARAGKIKGFTGIDDPYEAFLNCEIEIKVEDGVCP 267

QY 323 SPRAMAKQVLCYLEENGYL 341
Db 268 PPAEMAGQVVTYLEEKGF 286

RESULT 3
Q9FJX1
ID Q9FJX1 PRELIMINARY; PRT; 290 AA.
AC Q9FJX1;
```

```
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE ADENYLISULFATE KINASE-LIKE PROTEIN.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX MEDLINE=98403884; PubMed=9734815;
RA Kotani H., Nakamura Y., Sato S., Asamizu E., Kaneko T., Miyajima N.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. VI.
RT Sequence features of the regions of 1,367,185 bp covered by 19
RT physically assigned P1 and TAC clones.";
RL DNA Res. 5:203-216(1998).
DR EMBL; AB013390; BAB08460.1; -.
DR InterPro; IPR002891; APS_kinase.
DR Pfam; PF01583; APS_kinase; 1.
DR ProDom; PD002350; APS_kinase; 1.
KW Kinase
SQ SEQUENCE 290 AA; 32195 MW; BB7FB12990A6F6F8 CRC64;

Query Match 41.3%; Score 755; DB 10; Length 290;
Best Local Similarity 69.8%; Pred. No. 2.7e-54;
Matches 141; Conservative 28; Mismatches 33; Indels 0; Gaps 0;

QY 141 GKSTNLMHNCILIGSDRQKLLGQKGVVWITGLSGSGKSTLACALSRELHCRGHLTYVLD 200
Db 79 GQKNIVWHDPCVTKSDRQELIKQKGVITGLSGSGKSSLACALSRELHCRGHLTYVLD 138

QY 201 DGDNLRHGLNRDLSEKAEDRAENIRRVGEVAKLFADAGVICIASLISPYRRDRDACLPH 260
Db 139 DGDNLRHGLNRDLSEKAEDRAENIRRVGEVAKLFADAGVICIASLISPYRRDRDACLPH 198

QY 261 PHSNTEVFDLPLKICARPDKGLYKARTGKIKGFTGIDDPYPPPINGEIVKMKDEE 320
Db 199 PGDFIEVFMVDPVHLVCEARDPKGLYKLARAGKIKGFTGIDDPYEAFLNCEIEIKVEDGV 258

QY 321 CPSPKAMAKQVLCYLEENGYLQ 342
Db 259 SSSLCMADIVSVYLDQGYLK 280

RESULT 4
Q9SOR9
ID Q9SOR9 PRELIMINARY; PRT; 152 AA.
AC Q9SOR9;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE PUTATIVE ADENYLISULFATE KINASE, 3' PARTIAL (FRAGMENT).
GN Tll118.1.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.,
RA Ronning C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
RA Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RT "Arabidopsis thaliana chromosome III BAC Tll118 genomic sequence.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC011698; AAF05850.1; -.
DR InterPro; IPR002891; APS_kinase.
DR Pfam; PF01583; APS_kinase; 1.
DR ProDom; PD002350; APS_kinase; 1.
```

```
KW Kinase. 152 152
FT NON_TER 152 AA; 16974 MW; DBAC812B71D50DE5 CRC64;
SQ SEQUENCE 152 AA; 16974 MW; DBAC812B71D50DE5 CRC64;

Query Match 34.6%; Score 632; DB 10; Length 152;
Best Local Similarity 77.0%; Pred. No. 1.5e-44;
Matches 117; Conservative 20; Mismatches 15; Indels 0; Gaps 0;

Qy 137 MSNIGKSTNLIWHNCLIGQSDRQKLLGQKGCVVWITGLSGSKSTLACALSRHLHCRGHL 196
Dy 1 MSTVGNSTNIFWQESPICKTERQKLNQKGCVVWITGLSGSKSTLACALSRHLNRGKL 60
Qy 197 TVLDGDNLRHGLNDRSLFRAEDRAENIRRVGEVAKLFADAGVTCIASLISPYRRDRDAC 256
Dy 61 SYLDGDNLRHGLNDRSLFRAEDRAENIRRVGEVAKLFADAGVTCIASLISPYRRDRDAC 120
Qy 257 RALLPHSNFIEVFDLPKICEARDPKGLYKL 288
Dy 121 REMIQNSSPIEFVFMNLSQLCEARDPKGLYKL 152

RESULT 5
Q97MT8 PRELIMINARY; PRT; 200 AA.
AC Q97MT8;
DT 01-OCT-2001 (Tremblrel. 18, Created)
DT 01-OCT-2001 (Tremblrel. 18, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE ADENYLYLSULFATP KINASE.
GN CAC0103.
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1488;
RN SEQUENCE FROM N.A.
RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RA MEDLINE=21359325; PubMed=11466286;
RX Neelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA Gibson R., Lee H.M., Dubois J., Olu J., Hitti J., Wolf Y.I.,
RA Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
RA Bennett G.N., Koonin E.V., Smith D.R.;
RA "Genome sequence and comparative analysis of the solvent-producing
RT bacterium Clostridium acetobutylicum."
RL J. Bacteriol. 183:4823-4838(2001).
DR EMBL; AE007523; AAK78088.1; -.
DR InterPro; IPR002891; APS_kinase.
DR Pfam; PF01583; APS_kinase; 1.
DR ProDom; PD002350; APS_kinase; 1.
KW Kinase; Complete proteome.
SQ SEQUENCE 200 AA; 23001 MW; FE3A3CAE53D81C1F CRC64;

Query Match 31.3%; Score 571; DB 16; Length 200;
Best Local Similarity 54.9%; Pred. No. 2.3e-39;
Matches 113; Conservative 35; Mismatches 48; Indels 10; Gaps 3;

Qy 139 NICKSTNLIWHNCLIGQSDRQKLLGQKGCVVWITGLSGSKSTLACALSRHLHCRGHLTY 198
Dy 2 NNKSTNVVWQETKIKRQNEKMLKQKGAFLWFTGLSGSKSTVASALEKKLYEMGLTY 61
Qy 199 VLGDGDNLRHGLNDRSLFRAEDRAENIRRVGEVAKLFADAGVTCIASLISPYRRDRDACRA 258
Dy 62 LLDGDNLRHGLNDRSLFRAEDRAENIRRVGEVAKLFADAGVTCIASLISPYRRDRNARK 121
Qy 259 LLPHSNFIETFDLPKICEARDPKGLYKLARTGKIGFTGIDDPYPPINGEIVT---K 315
Dy 122 LL-GKDFVEVYIDCPIEVCEKRPQKGIYKARNGEIKNFTGIDSPYEKPEKPEITVETK 180
Qy 316 MKDECPSPKAMAKQVLCYLENGYL 341
Dy 181 DTEKC-----VDNILEYKQHKIL 200
```

```
RESULT 6
Q9NDP8 PRELIMINARY; PRT; 618 AA.
AC Q9NDP8;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE ATP SULFURYLASE/APS KINASE.
GN CI-ASAK.
OS Ciona intestinalis.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Phlebobranchia;
OC Cionidae; Ciona.
OX NCBI_TaxID=7719;
RN SEQUENCE FROM N.A.
RP SEQUENCE FROM N.A.
RA Hotta K., Takahashi H., Asakura T., Saitoh B., Takatori N., Satou Y.,
RA Satoh N.;
RT "Characterization of Brachyury downstream notochord genes in the Ciona
RT intestinalis embryo."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB036852; BAB00629.1; -.
DR InterPro; IPR002891; APS_kinase.
DR InterPro; IPR002650; ATP-sulfurylase.
DR Pfam; PF01583; APS_kinase; 1.
DR Pfam; PF01747; ATP-sulfurylase; 1.
DR ProDom; PD002350; APS_kinase; 1.
DR ProDom; PD002381; ATP-sulfurylase; 1.
KW Kinase.
SQ SEQUENCE 618 AA; 69559 MW; 75223CCB022BC297 CRC64;

Query Match 30.8%; Score 563; DB 5; Length 618;
Best Local Similarity 54.5%; Pred. No. 4.5e-38;
Matches 114; Conservative 30; Mismatches 51; Indels 14; Gaps 5;

Qy 122 APGEAPH-SPVKEKPYMSNIGKSTNLIWHNCLIGQSDRQKLLGQ---KGVVWITGLSGS 177
Dy 2 ASATANHGDDIQEKPPM-----SSNIVYQHHVSRDGRGVWGKGAFCGTVMFTGLSGA 56
Qy 178 GKSTLACALSRHLHCRGHLTYVLGDGDNLRHGLNDRSLFRAEDRAENIRRVGEVAKLFADA 237
Dy 57 GKSTLSMKLEIYCSGKIPAYSLDGDGNIHGLNDRSLFRAEDRAENIRRVGEVAKLFADA 116
Qy 238 GVICIASLISPYRRDRDACRALLPHSN--FIEVFIDLPLKICEARDPKGLYKLARTGKIK 295
Dy 117 GVCLVSVFISPYRKDRQSSARVHNKSNLPFIEVFVDTPLOVCEGRDVQKGLYKKGQGIK 176
Qy 296 GFTGIDDPYPPINGEIVTKMKD---EEC 321
Dy 177 GFTGLDSPYEPPEAEVYTKTADMPVEDC 205

RESULT 7
Q81351 PRELIMINARY; PRT; 271 AA.
AC Q81351;
DT 01-NOV-1998 (Tremblrel. 08, Created)
DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE 5'-ADENYLYLSULFATE KINASE (EC 2.7.1.25).
OS Enteromorpha intestinalis (Hollow green seaweed).
OC Eukaryota; Viridiplantae; Chlorophyta; Ulvophyceae; Ulvales; Ulvaceae;
OC Enteromorpha.
OX NCBI_TaxID=3116;
RN SEQUENCE FROM N.A.
RP SEQUENCE FROM N.A.
RA Gao Y., Leustek T.;
RT "Cloning of the cDNA for 5'-adenylylsulfate (APS) kinase from the
RT marine macroalga Enteromorpha intestinalis."
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF069952; AAC28856.1; -.
DR InterPro; IPR002891; APS_kinase.
DR Pfam; PF01583; APS_kinase; 1.
DR ProDom; PD002350; APS_kinase; 1.
```



```
RA Utterback T., Tran K., Wolf A., Vamathavan J., Ermolaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.:
RT "Complete genome sequence of Caulobacter crescentus.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
CC -I- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
DR EMBL: AE005842; AAK23461.1; -.
DR TIGR: CC1482; -.
DR InterPro: IPR002891; APS_kinase.
DR InterPro: IPR000795; GTP_EFTU.
DR InterPro: IPR004161; GTP_EFTU_D2.
DR Pfam: PF01583; APS_kinase.1.
DR Pfam: PF00009; GTP_EFTU; 1.
DR Pfam: PF03144; GTP_EFTU_D2; 1.
DR PRINTS: P00315; ELONGATNFC.
DR ProDom: PD002350; APS_kinase.1.
DR PROSITE: PS00301; EFATOR_GTP; UNKNOWN.1.
DR Complete proteome: GTP-binding; Transferrase.
KW SEQUENCE 635 AA; 69888 MW; 567FBBE2AE65345D CRC64;

Query Match 27.8%; Score 508.5; DB 16; Length 635;
Best Local Similarity 51.3%; Pred. No. 1.4e-33;
Matches 100; Conservative 32; Mismatches 60; Indels 3; Gaps 1;

QY 142 KSTNLIWHNCLIGSDRQKLLGQGVVWITGLSGSGKSTLACALSRELHCRHGLTYVLD 201
DB 436 RADNIHWQHTDVTKVSRAALKRQGVVWLTGLSGAGKSTIANLVKRLHALGRHTYLLD 495
QY 202 GDNLRHGLNRLSFKAEDRAENIRRVGEVAKLFADAGVICIASLISPYRRDRDACLALP 261
DB 496 GDNVRHGLNKDGLFTEEDRVENIRRVAEVAKLMDAGLIVITAFISPFRAERQLARDLE 555
QY 262 HSNFIEVFDLPLKICEARDPKGLYKARTGKIGFTGIDDPYEPINGEIVIKMKDEC 321
DB 556 PGFEIEVFVDTPVACARDVKGLYKARSQGLKNFTGVDSPYEAPESPRLRI---DTTA 612
QY 322 PSPKAMAKQVLCYLE 336
DB 613 IDPVEAAERIVAWLE 627

RESULT 11
ID Q22501 PRELIMINARY; PRT; 652 AA.
AC Q22501;
DT 01-NOV-1996 (TremBrel. 01, Created)
DT 01-NOV-1996 (TremBrel. 01, Last sequence update)
DE T14G10.1 PROTEIN.
GN T14G10.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]_
RP SEQUENCE FROM N.A.
RA Wild A.;
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
RS [2]
RR SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C. elegans. A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL: Z68860; CAA93098.1; -.
DR InterPro: IPR002891; APS_kinase.
DR InterPro: IPR002650; ATP_sulfurylase.
DR Pfam: PF01583; APS_kinase.1.
DR Pfam: PF01747; ATP_sulfurylase; 1.
DR ProDom: PD002350; APS_kinase; 1.
DR ProDom: PD002381; ATP_sulfurylase; 1.
SQ SEQUENCE 652 AA; 72987 MW; 36148F20C31AD875 CRC64;
```

```
Query Match 27.7%; Score 505; DB 5; Length 652;
Best Local Similarity 48.0%; Pred. No. 2.8e-33;
Matches 109; Conservative 38; Mismatches 64; Indels 16; Gaps 6;

QY 125 EAPHSPVKEKPVMSNI-GKSTNLIWHNCLIGSDRQKLLGQ---KGCVVWITGLSGSGK 179
DB 10 EGDAMPMLKKPRYSLSGQSTNITYQEHITISREAAAVGRHSGFRGCTIWFGLSGAGK 69
QY 180 STLACALSRELHCRHGLTYVLDGDNLRHGLNRLSFKAEDRAENIRRVGEVAKLFADAGV 239
DB 70 TTISFALERTLNKLGIPCTYGLDGNLRHGLNRLSFKAEDRAENIRRVGEVAKLFADAGV 129
QY 240 ICTASLISPYRRDRDACLALPHSN--FIEVFIDLPLKICEARDPKGLYKARTGKIGF 297
DB 130 ICIAAIFSPFQEDRLDARKIHESENVKFEIVHVTLEVEQDRDPKQLYKKARAGQILGF 189
QY 298 TGIDDPYEPINGEIVIKM-KD--EECPSPKAMAKQVLCYLEENGYL 341
DB 190 TGIDSAYEPPENAEIILDAGKGVQOC-----VQKVLHLESKGLL 230

RESULT 12
ID Q93HS6 PRELIMINARY; PRT; 627 AA.
AC Q93HS6;
DT 01-DEC-2001 (TremBrel. 19, Created)
DT 01-DEC-2001 (TremBrel. 19, Last sequence update)
DE NODO.
GN NODO.
OS Bradyrhizobium elkanii.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Bradyrhizobium group; Bradyrhizobium.
OX NCBI_TaxID=29448;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=USDA 94;
RX MEDLINE=21536155; PubMed=11679318;
RA Yasuta T., Okazaki S., Mitsui H., Yuhashi K., Ezura H., Minamisawa K.;
RT "DNA sequence and mutational analysis of rhizobitoxine biosynthesis
RT genes in Bradyrhizobium elkanii.";
RL Appl. Environ. Microbiol. 67:4999-5009(2001).
DR EMBL: AB062279; BAB55899.1; -.
SQ SEQUENCE 627 AA; 69085 MW; 6D2A74B9F9C7978 CRC64;
```

```
Query Match 27.6%; Score 504.5; DB 2; Length 627;
Best Local Similarity 49.7%; Pred. No. 3e-33;
Matches 95; Conservative 34; Mismatches 59; Indels 3; Gaps 1;

QY 142 KSTNLIWHNCLIGSDRQKLLGQGVVWITGLSGSGKSTLACALSRELHCRHGLTYVLD 201
DB 429 RATNVAPQPVAVGRKERADLNKQKPCIWFGLSGAGKSTIANIVQKLFAMSRHMLLD 488
QY 202 GDNLRHGLNRLSFKAEDRAENIRRVGEVAKLFADAGVICIASLISPYRRDRDACLALP 261
DB 489 GDNLRHGLNADLGFSEVDRMENIRRAEAAKLMDADSLIVICSFISPHRSEMDYRSLVS 548
QY 262 HSNFIEVFDLPLKICEARDPKGLYKARTGKIGFTGIDDPYEPINGEIVIKMKDEC 321
DB 549 KEFVEVFVDTPIECARRDPKGLYSKASKIKNFTGIDASAEAPIRPEIHLRTMEQ-- 606
QY 322 PSPKAMAKQVLC 332
DB 607 -TPEQMAQAVI 616

RESULT 13
ID Q9P7G9 PRELIMINARY; PRT; 202 AA.
AC Q9P7G9
DT 01-OCT-2000 (TremBrel. 15, Created)
DT 01-OCT-2000 (TremBrel. 15, Last sequence update)
DT 01-JUN-2001 (TremBrel. 17, Last annotation update)
```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 2, 2002, 01:52:23 ; Search time 2249 Seconds
(without alignments)
3191.553 Million cell updates/sec

Title: US-09-720-384A-4
Perfect score: 1826
Sequence: 1 RPPHFINTQTEPLVTHPTQPPP.....PKAMAKQVLCYLENGYLQA 343

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q/cgr2_1/USFTQ_spool/US09720384/runat_30102002_090527_19181/app_query.fasta_1.519
-DB=GenEmbl -OFMT=fastcap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human4.0.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09720384@cgn_1_1_1773_@runat_30102002_090527_19181 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*

29: em.vi.*
30: em.htg_hum.*
31: em.htg_inv.*
32: em.htg_other.*
33: em.htgo_inv.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	833	45.6	1086	8	AF044285	AF044285 Catharant
2	816	44.7	1077	8	ATU05238	U05238 Arabidopsis
3	816	44.7	1143	8	AY054287	AY054287 Arabidops
4	816	44.7	1185	8	ATCLAPSK	X75782 A.thaliana
5	815.5	44.7	970	8	AF178976	AF178976 Zea mays
6	778.5	42.6	1243	8	AF462823	AF462823 Arabidops
7	778.5	42.6	1311	8	AF043351	AF043351 Arabidops
8	633.5	34.7	3489	8	ATU59759	U59759 Arabidopsis
9	633.5	34.7	81298	8	AC004705	AC004705 Arabidops
10	632.5	34.6	51860	8	AB013390	AB013390 Arabidops
11	629.5	34.5	112067	8	ATCHR192	AL161596 Arabidops
12	629.5	34.5	118267	8	ATT5J17	AL1035708 Arabidops
13	595.5	32.6	1017	8	AF069952	AF069952 Enteromor
14	587.5	32.2	101410	8	ATAC009540	AC009540 Arabidops
15	577	31.6	11131	1	AE007523	AE007523 Clostridi
16	565.5	31.0	53533	1	BSY09476	Y09476 B.subtilis
17	565.5	31.0	210440	1	BSUB0006	Z99109 Bacillus su
18	563	30.8	2378	3	AB036852	AB036852 Ciona int
19	552	30.2	154330	2	AC104429	AC104429 Oryza sat
20	536	29.4	300550	1	AP001512	AP001512 Bacillus
21	533	29.2	11908	1	AE005822	AE005822 Caulobact
22	530	29.0	315000	1	RME603644	AL603644 Rhizobium
23	519	28.4	6386	1	AF308468	AF308468 Klebsiell
24	518.5	28.4	3282	1	RBU59507	U59507 Rhizobium B
25	514	28.1	28401	1	AB062279	AB062279 Bradyrhiz
26	513	28.1	977	8	PCU39393	U39393 Penicillium
27	510	27.9	220050	1	AJ141456	AJ141456 Yersinia
28	509.5	27.9	4042	1	RJU47272	U47272 Rhizobium t
29	509.5	27.9	6013	1	RPNDDHFG	X87608 Rhizobium m
30	503	27.5	3476	1	RHMDDPQA	X14809 Rhizobium m
31	503	27.5	4037	1	RMNDDPQ	AE007237 Sinoorhiz
32	503	27.5	14188	1	AE021701	AX281701 Sequence
33	502.5	27.5	2646	6	AX281701	L39001 Urechis cau
34	502.5	27.5	2819	3	UUNPASV	M94886 Azospirillum
35	500.5	27.4	3561	1	AZSNDDPQ	AF004875 Cavia por
36	498.5	27.3	2341	10	AF004875	AE003980 Xylella f
37	498.5	27.3	11117	1	AE003980	AE003980 Brucella
38	497.5	27.2	10580	1	AE003980	AF091242 Homo sapi
39	496	27.2	2014	9	AF091242	AF150754 Homo sapi
40	496	27.2	2377	9	AF150754	U34883 Mus musculu
41	496	27.2	2479	10	MMU34883	AR044160 Sequence
42	496	27.2	2506	6	AR044160	AF105227 Homo sapi
43	496	27.2	2537	9	AF105227	AF251798 Cavia por
44	496	27.2	3480	10	AF251798	E27970 APS kinase/
45	495	27.1	1875	6	E27970	

ALIGNMENTS

RESULT 1
AF044285
LOCUS AF044285
DEFINITION Catharantus roseus adenosine-5'-phosphosulfate-kinase (Craki)
mRNA, complete cds.
ACCESSION AF044285
VERSION AF044285.1
KEYWORDS GI:2832299
SOURCE Madagascar periwinkle.

```

ORGANISM Catharanthus roseus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asterales; euasterids I; Gentianales; Apocynaceae; Rauvolfioideae;
Vinceae; Catharanthus.
REFERENCE 1 (bases 1 to 1086)
AUTHORS Arz,H.E., Gisselmann,G., Schiffmann,S. and Schwenn,J.D.
TITLE A cDNA for adenylyl sulphate (APS)-kinase from Arabidopsis thaliana
JOURNAL Biochim. Biophys. Acta 1218 (3), 447-452 (1994)
MEDLINE 94325358
REFERENCE 2 (bases 1 to 1086)
AUTHORS Schiffmann,S. and Schwenn,J.D.
TITLE Isolation of cDNA clones encoding
adenosine-5'-phosphosulfate-kinase (EC2.7.1.25) from Catharanthus
roseus (Accession No. AF044285) and an isoform (akn2) from
Arabidopsis (Accession No. AF043351) (PGR98-116)
JOURNAL Plant Physiol. 117 (3), 1125 (1998)
REFERENCE 3 (bases 1 to 1086)
AUTHORS Schiffmann,S. and Schwenn,J.D.
TITLE Direct Submission
JOURNAL Submitted (23-JAN-1998) Biochemie der Pflanzen, Ruhr-Universitaet
Bochum, Universitaetsstrasse 150, Bochum 44780, Germany
LOCATION/Qualifiers
FEATURES
source
1..1086
/organism="Catharanthus roseus"
/db_xref="taxon:4058"
/sub_clone="KSCRakn"
/note="heterotrophic cell culture grown under sulfate
deficiency"
1..1086
/gene="Crakn"
39..977
/gene="Crakn"
/EC_number="2.7.1.25"
/function="phosphorylates 3'-OH group of
adenosine-5'-phosphosulfate"
/note="APS-kinase; putative processing site after residue
46"
/codon_start=1
/product="adenosine-5'-phosphosulfate-kinase"
/protein_id="AAC31145.1"
/db_xref="GI:2832300"
/translation="MIGSVKRPVSCVLPEDFTSTGLGKSSVKLPVNFAGSGG
GGVKGFLAPLKATEGSKTSFOVNGKVDNPHLOPDSNDSILNCGPQKK
ILQTTVGNSTNLHKKCAVEKSERQEPLOQRCVITWIGLSGSKSTLACALSRGLH
AKGLYTLIDGNVRLHNSDLSPKADRENIRIGVEAKLIFADAGVICIASLISPY
RKPPDACRLLPEGDFIEVMDVPLKVEARDPKGLYKLARAGKIKGFTGIDDPYEPP
LKSEIVLHOKLGMCDSPCDLADIVISYLENGYLKA"
BASE COUNT 313 a 187 c 267 g 319 t
ORIGIN

Alignment Scores:
Pred. No.: 9,18e-31 Length: 1086
Score: 833.00 Matches: 156
Percent Similarity: 84.11% Conservative: 24
Best Local Similarity: 72.90% Mismatches: 34
Query Match: 45.62% Indels: 0
DB: 8 Gaps: 0

US-09-720-384A-4 (1-343) x AF044285 (1-1086)
QY 130 ProValLysGluLysProValMetSerAsnIleGlyLysSerThrAsnIleLeuTrpHis 149
Db 333 CCAGGTAAAGATATTACAAACACTACCGTTGGAACTCTACTATATCTTTGGCAT 392
QY 150 AsnCysLeuIleGlyGlnSerAspArgGlnLysLeuGlyGlnLysGlyCysValVal 169
Db 393 AAATGTCTGTGGAAAAAGTGAAGGCGAGGAAACCCCTGCAGCAAGGGGCTGTGTATA 452
QY 170 TrpIleThrGlyLysSerGlyLysSerThrLeuAlaCysAlaLeuSerArgGlu 189
Db 453 TGGATTACTGCTCTAGTGTTCAGGAAGACGACCTTGGCATGTGCTTTGACTCGAGGC 512

```

```

QY 190 LeuHisCysArgGlyHisLeuThrTyValLeuAspGlyAspAsnLeuArgHisGlyLeu 209
Db 513 TTACATGCAAGGAAACTAACCTACATCTTGATGGTGATAATGTTTCGACATGGATTG 572
QY 210 AsnArgAspLeuSerPheLysAlaGluAspArgAlaGluAsnIleArgValGlyGlu 229
Db 573 AACAGTGCATCTTAGTCTTTAAAGCAGAAAGATCGAGCGGAGAAATATTAGGAGGATTGGAGAG 632
QY 230 ValAlaLysLeuPheAlaAspAlaGlyValIleCysIleAlaSerLeuIleSerProTyr 249
Db 633 GTTCAAAGCTCTTTCGTCGATGGAGTCATTTGCATTGCAAGTTTGATATCTCCCTAC 692
QY 250 ArgArgAspArgAspAlaCysArgAlaLeuLeuProHisSerAsnPheIleGluValPhe 269
Db 693 AGAAAGCCCCCGATGCGCTGCTTCTTCCAGAGGGGATTTTCATGAGGTATTC 732
QY 270 IleAspLeuProLeuLysIleCysGluAlaArgAspProLysGlyLeuTyrLysLeuAla 289
Db 753 ATGATGTCCTCTGAAAGTCTGCGAGCAAGGACCCCTAAGGCTTATATAAGCTTGCA 812
QY 290 ArgThrGlyLysIleLysGlyPheThrGlyIleAspAspProTyrGluProPheIleAsn 309
Db 813 CGAGCTGGTAAATCAAAAGGCTTTACTGGAATTGATGATCTTACGAGCCGCCACTAAA 872
QY 310 GlyGluIleValIleLysMetLysAspGluGluCysProSerProLysAlaMetAlaLys 329
Db 873 TCAGAGATTGTTTCATCAGAAAGCTTGGTATGCTGATTCCTCATGTGATTGGCTGAT 932
QY 330 GlnValLeuCysTyrLeuGluGluAsnGlyTyrLeuGlnAla 343
Db 933 ATAGTGATATCATATCTGGAGGAAATGGATATCTAAAGGCA 974

RESULT 2
ATU05238 1077 bp mRNA linear PLN 23-JUN-1998
LOCUS Arabidopsis thaliana APS kinase mRNA, complete cds.
DEFINITION U05238
ACCESSION U05238
VERSION U05238.1 GI:450234
KEYWORDS thale cress.
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE 1 (bases 1 to 1077)
AUTHORS Jain,A. and Leustek,T.
TITLE A cDNA clone for 5'-adenylylphosphosulfate kinase from Arabidopsis
thaliana
JOURNAL Plant Physiol. 105, 771-772 (1994)
MEDLINE 94345022
REFERENCE 2 (bases 1 to 1077)
AUTHORS Lee,S. and Leustek,T.
TITLE APS kinase from Arabidopsis thaliana: genomic organization,
expression, and kinetic analysis of the recombinant enzyme
JOURNAL Biochem. Biophys. Res. Commun. 247 (1), 171-175 (1998)
MEDLINE 98300303
REFERENCE 3 (bases 1 to 1077)
AUTHORS Leustek,T.
TITLE Direct Submission
JOURNAL Submitted (18-JAN-1994) Thomas Leustek, Center for Agricultural
Molec. Biology, Rutgers University, Cook College, College Farm
Road, New Brunswick, NJ 08903, USA
LOCATION/Qualifiers
1..1077
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/clone="ATK"
/tissue_type="whole plant"
33..863
/EC_number="2.7.1.25"
/functions="phosphorylates 5'-adenylylsulfate on
3'-hydroxyl group"
/note="similar to yeast APS kinase (MET14): Swissprot

```


5'UTR
CDS
/clone="RAFL09-19-L13(R13881)"
/note="ecotype: Columbia"
1..92
93..923
/note="putative adenosine phosphosulfate kinase"
/codon_start=1
/product="At2g14750/F26C24.11"
/protein_id="AAL06946.1"
/db_xref="GI:15810038"
/translation="MIAAGAKSLGLSMASPKGIFDSNMSNSRVSVVVACVSMDS
OTLSHNKSGSIPVKSINGHTGKQKQPLSTVGNSTNIKWHCSVEKVDRLDDKGC
VWVTLGSGSKSTLACALNMQYKGLCYLDGDNVNRHGLNDRDLSEKRAEDRAENIR
RVEVAKLFADAGIICIASLISPYTRDACSRLLPEDGFVEFMDVPLSVCEARDPK
GUYKLARAGKIKGFTGIDDDPYEPPLNCEISLREGGTSPIEMAEKVVGYLNDKGYLQA
924..1143

3'UTR
BASE COUNT 322 a 189 c 262 g 370 t
ORIGIN
Alignment Scores:
Pred. No.: 5,89e-30 Length: 1143
Score: 816.00 Matches: 159
Percent Similarity: 77.87% Conservative: 31
Best Local Similarity: 65.16% Mismatches: 44
Query Match: 44.69% Indels: 10
DB: 8 Gaps: 3

US-09-720-384a-4 (1-343) x AY054287 (1-1143)

QY 100 GlyAlaArgThrHisCysHisArgGlyIleGlyArgTrpValArgArgArgArgAsn 119
DB 219 GGATCTCAATCGATGATCAACAAATGATCTATCTCGAGGTAAATCCATTAAC 278
QY 120 GlyAlaAlaProGlyGluAlaProHisSerProValLysGluLysProValMetSerAsn 139
DB 279 GGT-----CACAGGGACAAAGCAAGACCT--TTGCTACG 314
QY 140 IleGlyLysSerThrAsnIleLeuTrpHisAsnCysLeuIleGlyLysSerAspArgLn 159
DB 315 GTCCGAAACTCGCAATATAAAGTGGCATGAATGTTCTTGAGAAAGTTGATAGACAG 374
QY 160 LysLeuLeuGlyLysGlyCysValValTrpIleThrGlyLeuSerGlySerGlyLys 179
DB 375 AGATTGCTTGATCAGAAAGGATGTGTGATTTGGGTCCACCGCTTCTAGTGTTACGGGAAG 434
QY 180 SerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTrpVal 199
DB 435 AGTATTGGCTTGTGTTGAATCAGATGTGTATCAAAAGGGGAGGCTTTGTTATATT 494
QY 200 LeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPhelLysAlaGluAsp 219
DB 495 CTTGATGCGTATATGTTAGCGATGCTTAAACCGTGATCTTAGCTTAAAGCTGAGGAT 554
QY 220 ArgAlaGluAsnIleArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyVal 239
DB 555 COTGCAGAGATATTCGTAGAGTTGGAGAGTTGCTAAGCTTTTTCGGATGCTGGATA 614
QY 240 IleCysIleAlaSerLeuIleSerProTrpArgArgAspAlaCysArgAlaLeu 259
DB 615 ATCTGATTCGAGTGTGATATCTCTTATAGACAGATAGGACGCTTGTCGAAAGTTTG 674
QY 260 LeuProHisSerAsnPhelIleGluValPheIleAspLeuProLeuLysIleCysGluAla 279
DB 675 CTCCTCCGAGGAGATTTGTTGAGTGTTCATGATGATGATGATGATGATGATGATG 734
QY 280 ArgAspProLysGlyLeuTrpLysLeuAlaArgThrGlyLysIleLysGlyPheThrGly 299
DB 735 AGGGATCAAAAGGCTCTTTACAAGCTTGCTGTCAGAGAGATCAAAAGGTTTATCCGGG 794
QY 300 IleAspAspProTrpGluProProIleAsnGlyGluIleValIleLysMetLysAspGlu 319
DB 795 ATCGATGACCTTTACGAGCCACCATTTGACTCCGAG-----ATTCTCTAGGACGTGA 848

QY 320 GluCysProSerProLysAlaMetAlaLysGlnValLeuCysTyrLeuGluGluAsnGly 339
DB 849 GGAGGAACCTTCTCTATCGAAATGGCGAAAGGTCCTCGGATACCTTAGATAACAAGGT 908
QY 340 TyrLeuGlnAla 343
DB 909 TATCTTCAAGCA 920
RESULT 4
ATCLAPSK 1185 bp mRNA linear PLN 05-FEB-1998
LOCUS A.thaliana (L.Heynh.) chloroplast mRNA for recombinant APS-kinase.
DEFINITION X75782
ACCESSION X75782.1 GI:414736
VERSION APS-kinase; ATP:adenylylsulfate-3'-phosphotransferase.
KEYWORDS thale cress.
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
REFERENCE 1 (bases 1 to 1185)
AUTHORS Arz.H.E., Gisselmann,G., Schiffmann,S. and Schwenn,J.D.
TITLE A cDNA for adenylyl sulphate (APS)-kinase from Arabidopsis thaliana
JOURNAL Biochim. Biophys. Acta 1218 (3), 447-452 (1994)
MEDLINE 94325358
REMARK (sites)
REFERENCE 2 (bases 1 to 1185)
AUTHORS Arz.H.E., Gisselmann,G., Schiffmann,S. and Schwenn,J.D.
TITLE A chloroplast APS-kinase cDNA from Arabidopsis thaliana
JOURNAL Unpublished
AUTHORS Schwenn,J.D.
TITLE Direct Submission
JOURNAL Submitted (03-NOV-1993) Schwenn J. D., Ruhr University Bochum,
Biology, Universitaetsstr. 150, 44780 Bochum, Germany
FEATURES
1..1185
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/clone="PM239x14cDNA"
/tissue_type="leaf"
/clone_lib="cDNA in lambda ZapII"
/dev_stage="vegetative"
42..872
/standard_name="ATP:adenylylsulfate-3'-phosphotransferase"
/EC_number="2.7.1.25"
/function="phosphorylates 3'-OH group of adenylylsulfate"
/citation=[2]
/codon_start=1
/product="APS kinase"
/protein_id="CAA53426.1"
/db_xref="GI:414737"
/db_xref="SPTREMBL:Q43295"
/translation="MIAAGAKSLGLSMASPKGIFDSNMSNSRVSVVVACVSMDS
OTLSHNKSGSIPVKSINGHTGKQKQPLSTVGNSTNIKWHCSVEKVDRLDDKGC
VWVTLGSGSKSTLACALNMQYKGLCYLDGDNVNRHGLNDRDLSEKRAEDRAENIR
RVEVAKLFADAGIICIASLISPYTRDACSRLLPEDGFVEFMDVPLSVCEARDPK
GUYKLARAGKIKGFTGIDDDPYEPPLNCEISLREGGTSPIEMAEKVVGYLNDKGYLQA
transit_peptide 42..155
/note="putative"
mat_peptide 162..869
/EC_number="2.7.1.25"
/note="putative mature peptide only - gene product
starting with ATC at position 162 is enzymatically active,
but maturation in the chloroplast has not yet been shown.
Homology to sequences from E. coli & ye"
/product="APS-kinase"
984..1012
985..990
stem_loop 1092..1120
polyA_site 267 g 383 t
stem_loop 335 a 200 c
BASE COUNT 335 a 200 c 267 g 383 t
ORIGIN

Alignment Scores:

Pred. No.: 6,11e-30 Length: 1185
Score: 816.00 Matches: 159
Percent Similarity: 77.87% Conservative: 31
Best Local Similarity: 65.16% Mismatches: 44
Query Match: 44.69% Indels: 10
DB: 8 Gaps: 3

US-09-720-384A-4 (1-343) x ATCLAPSK (1-1185)

QY 100 GlyAlaAArgThrHisCysHisArgGlyIleGlyArgTrpValArgArgArgArgAsn 119
Db 168 GGATCTCAAACTCGATGATACACAAAATGGATCTATTCCTGAGGTAAATCCATTAAAC 227
QY 120 GlyAlaAProGlyGluAlaProHisSerProVallysGluLysProValMetSerAsn 139
Db 228 GGT-----CACACGGGACAAAAGCAAGGACCT---TTGTCTACG 263
QY 140 IleGlyLysSerThrAsnIleLeuTrpHisAsnCysLeuIleGlyGlnSerAspArgGln 159
Db 264 GTCTGAAACTCGACAATATAAGTGGCATGAATGTTCTGTGAGAAAGTTGATAGACAG 323
QY 160 LysLeuLeuGlyGlnLysGlyCysValTrpIleThrGlyLeuSerGlySerGlyLys 179
Db 324 AGATTGCTGATCATCAAGAAGATGTGATTTGGGTACCGGCTCTAGTGTTCAGGGAAG 383
QY 180 SerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTyrVal 199
Db 384 AGTACTTTGGCTTGCTGTTTGAATCAGATGTTGTATCAAAAGGGAAGCTTTGTTATAT 443
QY 200 LeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLysAlaGluAsp 219
Db 444 CTGTAGTGTAATGTTAGGCATGCTTAAACCGTCATCTTACCTTTAAGCTGAGGAT 503
QY 220 ArgAlaGluAsnIleArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyVal 239
Db 504 CGTCAGAGAATATTCGTAGATGTGGAGAGGTGCTAAGCTTTTTCGGATGCTGAATA 563
QY 240 IleCysIleAlaSerLeuIleSerProTyrArgArgAspArgAlaCysArgAlaLeu 259
Db 564 ATCTCATTCGAGTTGATATCTCTATAGACAGATGAGCAGCTTGTCAAGATTG 623
QY 260 LeuProHisSerAsnPheIleGluValPheIleAspLeuProLeuLysIleCysGluAla 279
Db 624 CTCCCGAGGAGGATTTTGTGAGGTGTTCAATGATGTACCGCTTAGTGTTCGAGGCG 683
QY 280 ArgAspProLysGlyLeuTyrLysLeuAlaArgThrGlyLysIleLysGlyPheThrGly 299
Db 684 AGGATCCAAAGGCTCTTACAAGCTTCTGTCGAGGAAGATCAAGGTTTACCGGG 743
QY 300 IleAspAspProTyrGluProProIleAsnGlyGluIleValIleLysMetLysAspGlu 319
Db 744 ATCGATGACCTTACGAGCCACCATTTGAACCTGCGAG-----ATTCTCTAGGAGTGAA 797
QY 320 GluCysProSerProLysAlaMetAlaLysGlnValLeuLysCysTyrLeuGluGluAsnGly 339
Db 798 GGAGAACTTCTCTCATGAATGCGGAAAGGTCGTCCGATACITTAGATACACAGGT 857
QY 340 TyrLeuGlnAla 343
Db 858 TATCTTCAAGCA 869

RESULT 5
AF178976
LOCUS AF178976 970 bp mRNA linear PLN 13-DEC-1999
DEFINITION Zea mays adenosine-5'-phosphosulfate kinase (AK1) mRNA, partial cds.

ACCESSION AF178976
VERSION AF178976.1 GI:6563284
KEYWORDS Zea mays.
SOURCE Zea mays.
ORGANISM Zea mays

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

/organism="Zea mays"

/cultivar="Dekalb XL72"

/db_xref="taxon:4577"

/gene="AK1"

/FC_number="2, 7, 1, 25"

/note="APS kinase; ATP adenosine-5'-phosphosulfate

3'-phosphotransferase"

/product="adenosine-5'-phosphosulfate kinase"

/protein_id="AAFL7236.1"

/db_xref="GI:6563285"

/translation="VPVPLLLAVARASRRRAPEPRVKSRVASAPPPPTSGMWNQ

NEGRRLKRGTPLVCTGDSRVEQHPAGVEGKAMKMSSTVPKSNIFWHDGCPVGK

TDRONVLKQKCVWLTGLSGSKSLACTLIGRELHTRGKLAVLDEGDLRLGNKDL

GFKAEADRENIRRVAKLFLADAGLVCIALSLSPRRDRSCRLDSSDFIEVFLNMS

LELCEARDPKGLKRLARAGKINGFTGIDDPYEAPLNCIEIKEIVGDCVCPPEMAGQV

VTYLEEGFLHE"

BASE COUNT 251 a 221 c 271 g 227 t
ORIGIN

Alignment Scores:

Pred. No.: 5, 27e-30 Length: 970
Score: 815.50 Matches: 178
Percent Similarity: 67.97% Conservative: 30
Best Local Similarity: 58.17% Mismatches: 65
Query Match: 44.66% Indels: 33
DB: 8 Gaps: 6

US-09-720-384A-4 (1-343) x AFL78976 (1-970)

QY 40 ProThrProThrLeuAlaValIleLeuValAsnProGlnArgAlaProProValLeuPro 59
Db 33 CCGCGCGCTCGCGTCCGCGACCTG-----AACCGGAGCCCGCGTCAAAATCTA 80
QY 60 GlyLeuThrProSerAspAlaProLeuProAlaLeuValIleHisGlyLeuThrProArg 79
Db 81 GGGTTGATCCGCGCGCGCCCTCTCCCGAGACTCCAGCGCTGGAACACCAAGACGAG 140
QY 80 SerSerHisSerSerAlaClyLeuAlaSerAspSerClyArgArgGluGlyGluClyArg 99
Db 141 GTG-----ACAGGCGACAAAGCTGCGAGGAGCG 170
QY 100 GlyAla-ArgThrHisCysHisArgGlyIleGlyArgTrpValArgArgArgAs 119
Db 171 CCCTCGTCAATGCAC-----TGCGGACCGCTTCCGTCGA----- 204
QY 119 nGlyAlaAProGlyGluAla-ProHisSerProVallysGluLys-----ProValM 137
Db 205 -----GGAGCAGCGCGAGCAGCAGGAGTGGGGAAAGAGCTTGGAAAATGT 251
QY 137 etSerAsnIleGlyLysSerThrAsnIleLeuTrpHisAsnCysLeuIleGlyGlnSerA 157
Db 252 CATCCACCGTGGCGAAGTCAATCAATATCTCTGGCACGATTGTCCAGTTGCAAGACTG 311
QY 157 spArgGlnLysLeuLeuGlyGlnLysGlyCysValValTrpIleThrGlyLeuSerGlyS 177

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 970)

Bolchi A., Petrucco, S. and Ottanello, S.

Isolation and comparative expression analysis of a maize cDNA

encoding adenosine 5'-phosphosulfate kinase

Unpublished

2 (bases 1 to 970)

Bolchi A., Petrucco, S. and Ottanello, S.

Direct Submission

Submitted (19-AUG-1999), Istituto di Scienze Biochimiche,

Universita' di Parma, Viale delle Scienze, Parma 43100, Italy

Location/Qualifiers

1, 970

/organism="Zea mays"

/cultivar="Dekalb XL72"

/db_xref="taxon:4577"

/gene="AK1"

/FC_number="2, 7, 1, 25"

/note="APS kinase; ATP adenosine-5'-phosphosulfate

3'-phosphotransferase"

/product="adenosine-5'-phosphosulfate kinase"

/protein_id="AAFL7236.1"

/db_xref="GI:6563285"

/translation="VPVPLLLAVARASRRRAPEPRVKSRVASAPPPPTSGMWNQ

NEGRRLKRGTPLVCTGDSRVEQHPAGVEGKAMKMSSTVPKSNIFWHDGCPVGK

TDRONVLKQKCVWLTGLSGSKSLACTLIGRELHTRGKLAVLDEGDLRLGNKDL

GFKAEADRENIRRVAKLFLADAGLVCIALSLSPRRDRSCRLDSSDFIEVFLNMS

LELCEARDPKGLKRLARAGKINGFTGIDDPYEAPLNCIEIKEIVGDCVCPPEMAGQV

VTYLEEGFLHE"

Db	282	TGATTTCCGCCCTTTCAGAACTATACTGAAAGACGGGAAGAAATAACCGA-----	333	
Qy	123	oGlyluAlaProHisSerProValYsGluLysProValMetSerAsnIleGlyLysSe	143	
Db	334	-----AAAGAG-----	AAAGC	344
Qy	143	rThraSnlLeuTrpHisAsnCysLeuIleGlyGlnSerAspArgGlnLysLeuLeuG1	163	
Db	345	AGAGAACATCGTGTGCACGACAGTTCGATATGCGAGATGCGACAGACACACATCTTCTCA	404	
Qy	163	yGlnLysGlyCysValItrPileThrGlyLeuSerGlySerGlyLysSerThrLeuAl	183	
Db	405	ACAAAGGGTGTGCTGTTGGATCACTGGTCTCAGTGGTTCAGGGGAAAGCACGTGTTC	464	
Qy	183	acysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTyrValLeuAspGlyAs	203	
Db	465	FTTGTCACATAGTAAGCATTTGTTCAAGAGGCAAACTTACTTACACACTCGACGGCGA	524	
Qy	203	pAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLysAlaGluAspArgAlaGluAs	223	
Db	525	CAATGTCGTCACGGCTTAACCGTACGCTCACTTTCAAAGCAGACGACCGACGAAAA	584	
Qy	223	nIleArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyValIleCysIleAl	243	
Db	585	CATTAGAGAAATTGGTGGTGGCTAGTGTGTTGCTGAGTCGGAGTCATTGTATAGC	644	
Qy	243	sSerLeuIleSerProTyrArgAspArgAspAlaCysArgAlaLeuLeuProHisSe	263	
Db	645	AAAGTTGATTTCTCCGTACCGGAGACAGACGCGTGCCTCTCTTCTTACCTGACGG	704	
Qy	263	rAsnPheIleGluValPheIleAspLeuProLeuLysIleCysGluAlaArgAspProLy	283	
Db	705	CGATTTCGTCAGGCTCTTACGACGTCCTTCATGTCGTGCGATCGAGATCCAAAA	764	
Qy	283	sGlyLeuTyrLysLeuAlaArgThrGlyLysIleGlySerGlyPheThrGlyIleAspPr	303	
Db	765	GGGGTGTACAAAGTTGCAGCTGCAGCAAAATCAAAAGGTTCACTGGAATCGACGACC	824	
Qy	303	oTyrGluProPheLysGlyGluIleValIleLys---MetLysAspGluGlyCysPr	322	
Db	825	TTACGAGCCACCATGTAATGCGAGTAGTGTCTGAACACACAGGAGACGAGCTCGTG	884	
Qy	322	oSerProLysAlaMetAlaLysGlnValLeuCysTyrLeuGluGluAsnGlyTyrLeuG1	342	
Db	885	TTGCGCAGTCAGATGCTGAGACATCATCTTACCTGCAAAACAAAGGTATCTTGA	944	
Qy	342	n	342	
Db	945	G	945	
RESULT 7				
AF043351				
LOCUS				
DEFINITION				
Arabidopsis thaliana adenosine-5'-phosphosulfate-kinase (akn2)				
mRNA, complete cds.				
ACCESSION				
AF043351.1				
VERSION				
GI:2829132				
KEYWORDS				
SOURCE				
thale cress.				
ORGANISM				
Arabidopsis thaliana				
AUTHORS				
Arz,H.E., Gisselmann,G., Schiffmann,S. and Schwenn,J.D.				
TITLE				
A cDNA for adenylyl sulphate (APS)-kinase from Arabidopsis thaliana				
JOURNAL				
Biochim. Biophys. Acta 1218 (3), 447-452 (1994)				
MEDLINE				
9432538				
REFERENCE				
2 (bases 1 to 1311)				
Schiffmann,S. and Schwenn,J.D.				
AUTHORS				
Isolation of cDNA clones encoding				
adenosine-5'-phosphosulfate-kinase (EC2.7.1.25) from Catharanthus				
roseus (Accession No. AF044285) and an isoform (akn2) from				
Arabidopsis (Accession No. AF043351) (FGR98-116)				
Plant Physiol. 117 (3), 1125 (1998)				
REFERENCE				
3 (bases 1 to 1311)				
AUTHORS				
Schiffmann,S. and Schwenn,J.D.				
TITLE				
Direct Submission				
JOURNAL				
Submitted (16-JAN-1998) Biochemie der Pflanzen, Ruhr-Universitaet				
Bochum, Universitaetsstrasse 150, Bochum 44780, Germany				
FEATURES				
source				
1..1311				
/organism="Arabidopsis thaliana"				
/cultivar="Columbia"				
/db_xref="taxon:3702"				
/clone="pGAD10"				
/tissue_type="green vegetative"				
/dev_stage="3 week old"				
1..1311				
/gene="akn2"				
61..942				
/gene="akn2"				
/EC_number="2.7.1.25"				
/function="phosphorylates 3'-OH group of				
adenosine-5'-phosphosulfate"				
/note="APS-kinase isoform; similar to the product encoded				
by GenBank Accession Number X75782"				
/codon_start=1				
/product="adenosine-5'-phosphosulfate-kinase"				
/protein_id="AAC39520.1"				
/db_xref="GI:2829133"				
/translation="MEGLAIRASRPVSFCISPLGSDSHRKPSPDGLKLPASSIPAD				
SRKLVANSTSFHPISAVNVSAOASLTADFPALSETILKEKNNNGKEAKENIVHESLI				
CRCDROOLQKCVVMITGLSGSKSTVACALSKALFERGKLTLYTLDGNDVHGLNR				
DLTFKASHRTENIRRIAGEVAKLFADVGVICIASLISPYRRDRDACSLLPDGDFEVE				
MDVPLHVCESRDKGLIKLARGKINGFTGIDDPYEAIVNCEVVLKHTGDDSCSPRO				
MAENLIYLNQNGYLEG"				
misc_feature				
238..240				
/gene="akn2"				
/note="putative processing site 3' of this position;				
unclassified site"				
BASE COUNT				
373 a 293 c 341 t 1 others				
ORIGIN				
Alignment Scores:				
Pred. No.: 3.64e-28 Length: 1311				
Score: 778.50 Matches: 175				
Percent Similarity: 66.04% Conservative: 37				
Best Local Similarity: 54.52% Mismatches: 67				
Query Match: 42.63% Indels: 43				
DB: 8 Gaps: 9				
US-09-720-384A-4 (1-343) x AF043351 (1-1311)				
Qy	40	ProThrProThrLeuAlaValIleLeuValAsnProGlnArgAlaProProVal-----	57	
Db	49	CCGTCGATTCATGAAGGATTAGCTATCAGACATCTCGA---CCGTCGGTTTCTGT	105	
Qy	58	---LeuProGlyLeu-----ThrProSerAspAla-----Pro	67	
Db	106	TCATTCAGGTCGCGCGGATTCACCAAGAACTCCAGGACGCTTCTCTCAAG	165	
Qy	68	LeuProAlaLeuValIleHisGlyLeuThrProArgSerSerHisSerSerAlaGlyLeu	87	
Db	166	CTGCCTCGCTC-GTCTATTCGCGGAGACGCGGAAATAGTAGGGAATTTACTTCTT	224	
Qy	88	AlaSerAspSerGlyArgGluGlyGluGlyAlaArgThrHisCysHisArg	107	
Db	225	TCATCCAAATCTCAGCCG-----TAACGTCCTGCTCAAGCTCCCTCACC	272	
Qy	108	GlyIleGlyArg-----TirPValArgArg-ArgArgArgAsnGlyAlaAlaPr	123	
Db	273	TGATTTTCCCGCCCTTTCAGAAACTATACTCAAAGAGGAAGAAATAACGA-----	324	
Qy	123	oGlyGluAlaProHisSerProValLysGluLysProValMetSerAsnIleGlyLysSe	143	

Db	325	-----AAAGAG-----	AAAGC	335
QY	143	rThrAsnIleLeuTrpHisAsnCysLeuIleGlyGlnSerAspArgGlnLysLeuG1	163	
Db	336	AGAACATCGTGGGCACAGAGTTCATATGCAGATGGCAGACAACTTCCTCA	395	
QY	163	yGlnGlyGlyCysValValTrpIleThrGlyLeuSerGlySerGlyLysSerThrLeuAl	183	
Db	396	ACAAAGAGGTGTGCTGTTGGATCACTGGTCTCAGTGGTTCAGGGAAAAAGCACTGTTC	455	
QY	183	acysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTyValLeuAspGlyAs	203	
Db	456	TTGTCCACTAGTAACCATTTGTTCAAGAGGCAACTTACTTACCACTCGACGGCA	515	
QY	203	pAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLysAlaGluAspArgAlaGluAs	223	
Db	516	CAATTCGTCACGGCTTAACCGTGAGCTCACTTCAAAAGCAGAGCACCGCAGAAA	575	
QY	223	nileArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyValIleCysIleAl	243	
Db	576	CATTAGAAGAAATGGTGAGTGGCTAAAGTTGTTGCTGACGTCGGAGTCATTGTATAGC	635	
QY	243	aSerLeuIleSerProTyrrArgArgAspAlaCysArgAlaLeuLeuProHisSe	263	
Db	636	AAAGTTGATTTCTCCGTACCGGAGACAGACGCGTCCGCTCTTGTACCTGACGG	695	
QY	263	rAsnPheIleGluValPheIleAspLeuProLeuLysIleCysGluAlaArgAspProLy	283	
Db	696	CGATTTCTCGAGGTCTTCATGACGTTCCCTTCATGCTGCGAGTCGAGATCCAAA	755	
QY	283	sglyLeuTyrrLysLeuAlaArgThrGlyLysIleLysGlyPheThrGlyIleAspPr	303	
Db	756	GGGGTTGTACAAAGTTGCGAGTCGAGGCAAAATCAAGGGCTTCACGGAATCGAGACCC	815	
QY	303	oTyrrGluProPheIleGlyValIleValIleLys---MetLysAspGluGluCysPr	322	
Db	816	TTACGAGCACCACTGAATGCGAGGTAGTGTCTGAAACACACAGAGACGAGTCTGTG	875	
QY	322	oSerProLysAlaMetAlaLysGlnValLeuCysTyrrLeuGluGluAsnGlyTyrrLeuG1	342	
Db	876	TTCCGACGTCAGATGGCTGAGAACATCATCTCTTACCTGCAAAAGGTTATCTTGA	935	
QY	342	n 342		
Db	936	G 936		
RESULT 8				
ATU59759				
LOCUS				
DEFINITION		3489 bp DNA linear	PLN 19-JUN-1998	
ACCESSION		Arabidopsis thaliana AFS Kinase gene, complete cds.		
VERSION		U59759.1		
KEYWORDS		thale cress.		
SOURCE		Arabidopsis thaliana		
ORGANISM		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.		
REFERENCE		1 (bases 1 to 3489)		
AUTHORS		Jain,A. and Leustek,T.		
TITLE		A cDNA clone for 5'-adenylylphosphosulfate kinase from Arabidopsis thaliana		
JOURNAL		Plant Physiol. 105 (2), 771-772 (1994)		
MEDLINE		94345022		
REFERENCE		2 (bases 1 to 3489)		
AUTHORS		Lee,S. and Leustek,T.		
TITLE		AFS Kinase from Arabidopsis thaliana: genomic organization, expression, and kinetic analysis of the recombinant enzyme		
JOURNAL		Biochem. Biophys. Res. Commun. 247 (1), 171-175 (1998)		
MEDLINE		98300303		
REFERENCE		3 (bases 1 to 3489)		
AUTHORS		Leustek,T.		
TITLE		Direct Submission		
JOURNAL				
FEATURES				
source				
Location/Qualifiers				
1..3489				
/organism="Arabidopsis thaliana"				
/strain="Lansberg erecta"				
/db_xref="taxon:3702"				
1486..1713				
join(1486..1713,1793..1933,2018..2167,2253..2371,2460..2544,2630..2676,2799..3105)				
join(1518..1713,1793..1933,2018..2167,2253..2371,2460..2544,2630..2676,2799..2891)				
/EC_number="2.7.1.25"				
/note="adenosine 5'-phosphosulfate kinase;				
5'-adenylylphosphosulfate kinase"				
/codon_start=1				
/product="AFS kinase"				
/protein_id="AAC50034.1"				
/db_xref="GI:1575322"				
/translation="MIAAGAKSLGLSWASPKGIFDSNSMSRSVVVVRACVSMDS				
OTLSHNKSGSIPEVKSINGHCGKQGPLSVGNSTATKHGECSEVKVDRORLLQKGC				
VIVWTGLSSGSKSTLACALNQMLVQKGLCTIILDGNVRHGLNRDLSPKADRAENIR				
RVGEVAKLFADAGICIAISLISPTIRDRACSLPPEGDFVEFMDVPLVSCEARDPK				
GLYKLARAGKIKGTGIDDPYEPPLNCEIISLUREGGTSPLEMAEKVVGVYLDNKGYLQA				
"				
1793..1933				
exon				
2018..2167				
exon				
2253..2371				
exon				
2460..2544				
exon				
2630..2676				
exon				
2799..3105				
exon				
BASE COUNT	1041 a	547 c	603 g	1298 t
ORIGIN				
Alignment Scores:				
Pred. No.:	4.8e-21	Length:	3489	
Score:	633.50	Matches:	152	
Percent Similarity:	49.05%	Conservative:	28	
Best Local Similarity:	41.42%	Mismatches:	31	
Query Match:	34.69%	Indels:	157	
DB:	8	Gaps:	6	
US-09-720-384A-4 (1-343) x ATU59759 (1-3489)				
QY	132	LysGluLysProValMetSerAsnIleGlyLysSerThrAsnIleLeuTrpHisAsnCys	151	
Db	1798	AAGCAAGGACCT---TTGCTACGGTCGGAACATGACAAATATAAAGTGCATGATCT	1854	
QY	152	LeuIleGlyGlnSerAspArgGlnLysLeuLeuGlyGlnLysGlyCysValValTrpIle	171	
Db	1855	TCTGTTGAGAAAGTTGATAGACAGAGATGCTTGATCAGAAAGGATGTGATTGGGTC	1914	
QY	172	ThrGlyLeuSerGlySer-----	177	
Db	1915	ACCGGCTCTAGTGGTTCAGGTTTGTGTGAACAAGCTTCAAGTTTGTGCTCTA	1974	
QY	178	-----		
Db	1975	TGTGAAATCTATCTAAATGTTCTCTATGATTTATAAATTTGTAGGAAGAGTACTTTGGCT	2034	
QY	184	CysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTyrrValLeuAspGlyAsp	203	
Db	2035	TGTGCTTTGAATCAGATGTTGTATCAAAAGGGAAGCTTTGTTATATTCTTGATGCTGAT	2094	
QY	204	AsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLysAlaGluAspArgAlaGluAsn	223	
Db	2095	AATGTTAGGATGCTTAAACCGTGATCTTAGCTTTAAAGCTGAGATCGTCAGAGAAT	2154	
QY	224	IleArgArgVal-----	227	
Db	2155	ATTCGTAGAGTTGGTGAGCTTTAATCTCTTTAATCCAATATTGTTGTTGTTTACGCTT	2214	
QY	228	-----GlyGluValAlaLysLeuPheAl	235	

QY	269	---PheIleAspLeuProLeuLysIleCysGluAlaArgAspProLysGlyLeuTyrLys	287
Db	19478	GTATTATTGATGGCACTCCATCTTTGTGAAGCTAGATGCCAAAGGCTTATACAAA	19537
QY	288	LeuAlaArgThrGlyLysIleLys-----	295
Db	19538	CGTCAGCGCGCTGGTAATCAAGTAACCTTTGCTTTCCTGCACTGTTTTCRACT	19597
QY	295	-----	295
Db	19598	TAACCCATGTTCTACTGATTTTCATTAAATTATATAACCTTCTTGATCATCACTTGATT	19657
QY	296	-----GlyPheThrGlyIleAspAsp	302
Db	19658	ATATGTAATATATTCTCCCAAAACATTTCTTCCTCCAGGTTTACAGGAGTAGATGAT	19717
QY	303	ProTyrGluProProIleAsnGlyLuleValIle-----	314
Db	19718	CCATATGAAGCGCTTTGGATTCCGAGGTTTCATATAATATCCAAATTTTGTATATAAGT	19777
QY	314	-----	314
Db	19778	TGGAATCTAGTCATCTCCGAACCTGTAATCTAATGTTGTGTTCCGACTACTCTTTGCAG	19837
QY	315	-----LysMetIlyAspGluCys	321
Db	19838	ATTGTAATACAAACAGTAGAGAACAGGGCGCTTCTTCTTCATCTTCATCTTCATCTTCA	19897
QY	322	ProSerProLysAla-----MetAlaLysGlnValLeuCystyrLeuGlu	336
Db	19898	CCTTCATCTTCGCTCTTCTCTCTCTGTTGTAATGGCAGATATTCTGTGTCGTA	19957
QY	337	GluAsnGlyTyrLeuGln	342
Db	19958	CAAAATGATACCTTGAAG	19975
RESULT 11			
ATCHRIV92			
LOCUS	ATCHRIV92	112067 bp	DNA linear PLN 16-MAR-2000
DEFINITION	Arabidopsis thaliana	DNA chromosome 4, contig fragment No. 92.	
ACCESSION	AL161596		
VERSION	AL161596.2	GI:7271037	
KEYWORDS	thale cress.		
SOURCE	Arabidopsis thaliana		
ORGANISM	Arabidopsis thaliana		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
AUTHORS	Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.		
JOURNAL	1 (bases 9546 to 9977)		
REFERENCE	Volckaert,G., Grymonprez,B., Voet,M., Robben,J., Mewes,H.W.,		
AUTHORS	Lemcke,K. and Mayer,K.F.X.		
JOURNAL	Unpublished		
REFERENCE	2 (bases 107578 to 112067)		
AUTHORS	Rose,M., Hempel,S., Entian,K.-D., Mewes,H.W., Lemcke,K. and		
JOURNAL	Mayer,K.F.X.		
REFERENCE	3 (bases 1 to 104339; 104340 to 107700)		
AUTHORS	Murphy,G., Ridley,P., Hudson,S., Mewes,H.W., Lemcke,K. and		
JOURNAL	Mayer,K.F.X.		
REFERENCE	Unpublished		
AUTHORS	4 (bases 1 to 112067)		
JOURNAL	EU Arabidopsis sequencing project.		
REFERENCE	Submitted (10-MAR-2000) MIPS, at the Max-Planck-Institut fuer		
AUTHORS	lence@mips.biochem.mpg.de, mayer@mips.biochem.mpg.de Project		
JOURNAL	Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge		
REFERENCE	Laboratory, John Innes Centre, Colney Lane, NR4 7UU Norwich, UK,		
AUTHORS	E-mail: michael.bevan@bbsrc.ac.uk		
JOURNAL	Information on performance of analysis and a more detailed		
REFERENCE	annotation of this entry and other sequences of chromosomes 3, 4		
AUTHORS	and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/thal/		
JOURNAL			
COMMENT			

this fragment has an overlap with ATCHRIV91 at the 5' end and an overlap with ATCHRIV93 at the 3' end.

Location/Qualifiers

1. 112067

/organism="Arabidopsis thaliana"

/variety="Columbia"

/db_xref="taxon:3702"

/chromosome="4"

6122..6415

/gene="AT4g39830"

/number=1

join(6122..6415,6481..6573,6652..6864,6959..7153,

7250..8203)

/gene="AT4g39830"

/note="strong similarity to L-ascorbate oxidase, Cucumis

sativus, PIR1:KSKVAO

Contains Glycosyl hydrolases family 5 signature

AA384-393;Multicopper oxidases signatures

AA131-151;Multicopper oxidases signatures

AA540-560;Multicopper oxidases signatures

AA545-556;Prokaryotic membrane lipoprotein lipid

attachment site AA567-577

contains EST gb:AI994215.1, AA585817, AI996008.1"

/codon_start=1

/product="putative L-ascorbate oxidase"

/protein_id="CAB80646.1"

/db_xref="GI:7271038"

/translation="MMRPASDSPTVHVNLMLVLCFTALFFSSVLCQGLRRFKVEVKY

EKSPDCFEKLVITNGEPFGTIKAQGGDTIVVLEKNSFMTEVAVHWHGIRLQIGTP

WFDGEGVTCQIPILGEVFIQTVVDRQGTIVVHSHYGMORESLGIMIQVSPATEP

EPFTYDNRFLNLDWYHKMSERATGLASIPFKWGEQSLMIOGRFNCNNLT

PPSLVSGVCNVSNADCSRILTVIPGKTYRLRGLSLTALSLSFQTEGHNTLVVYADG

HYVEPTVKLVFVYSGEYVLLKADQPNRYNTSSIVSRPATTTPATVLYTPN

HPRRPPTSESSIVPEWNTSRSLAQSIAIKARGFHALPENSCKVILVLLNTQNEV

NGYRMSYNNVSYHHPKTPYLIALKQNLTAEDRTPAPENYDSRYDIPAKPLNANA

TTSDGIYRLRFNSTVDVILQNTANNMNSSETHPWHLHGHDFWVLYGEGKFNSEDP

KRYNRVDPIKKNTAVAPGFWTALFRADNPVWSPFCHIESHFFMGMGIVFESGIDK

VSSLPSISSIMCGCGQTKR"

6122..8203

/gene="AT4g39830"

6416..6480

/gene="AT4g39830"

/number=1

6481..6573

/gene="AT4g39830"

/number=2

6574..6651

/gene="AT4g39830"

/number=2

6652..6864

/gene="AT4g39830"

/number=3

6865..6958

/gene="AT4g39830"

/number=3

6959..7153

/gene="AT4g39830"

/number=4

7154..7249

/gene="AT4g39830"

/number=4

7250..8203

/gene="AT4g39830"

/number=5

12287..13642

/gene="AT4g39840"

12287..13642

/gene="AT4g39840"

/note="contains EST gb:AI999725.1, Z46537, F13826,

AI996686.1"

/codon_start=1

/product="putative protein"

/protein_id="CAB80647.1"

```
/db_xref="GI:7271039"
/translation="MSPLKYLCPFRKVKLLLLLYSSLFSSSSFFASSVDDEHQLQRIDP
HFVRRLVKDDSDDDDETNPPLPPKKKILLTDSVSSSSSGTKKNGTKLLRPIS
SSSTKNQKAKTKTTIKPPSGKSLNKKSTKSNNTTKTSELKKNLNGSTKSTNSTSSIKKS
ADLSKSSKNKTTIKPPSKSLSPSPSKSQSPKVPKTKSQSEKIKPFWLDDDEED
EDVSFRLPTRFRQSLIPDLERSITGSKNINKANKOITKNEKPYFGNKYAPTIAS
VVSFVLPVLLVSLIFNRFKAFVSLQKILFIOIYLSIVFSICLSSLVGTIEPLK
FLYATSSPYVCLQILOTGLGVVFLVLLLMYLVLFVSDGGLGKVLGLAQTPEGFAV
GLHYVAVFHRVLRQPPKTNKIHGYATCFLLICLLSNAERRKKEYLEGGDEGKK
N"
exon      12287..13642
           /gene="AT4g39840"
           /number=1
misc_feature 1539..16205
           /note="EST N95877 matches to positions 26106-26355,
           26723-26772; possibly related to subsequent gen"
exon      16239..16300
           /gene="AT4g39850"
           /number=1
gene      16239..23312
           /gene="AT4g39850"
CDS       join(16239..16300,16394..16679,16981..17205,17341..17538,
           17675..18129,18222..18357,18433..18511,19039..19271,
           19326..19630,19712..19844,19901..20056,20213..20435,
           20515..20623,20845..20935,21205..21372,21453..21568,
           21626..21860,21924..22104,22162..22362,22534..22680,
           22750..22983,23134..23312)
           /gene="AT4g39850"
           /note="similarity to 70K peroxisomal membrane protein,
           Rattus norvegicus , PIR2:A35723
           Contains ABC transporters family signature AA580-594; ABC
           transporters family signature AA1369-1383; ATP/GTP-binding
           site motif A (P-loop) AA1247-1254
           contains EST gb:N95877"
           /codon_start=1
           /product="putative protein"
           /protein_id="CAB80648.1"
           /db_xref="GI:7271040"
           /translation="MPSLQLQLTERGRGLVSRKRSILLAAIGIVAAGTNAVYLSRV
           KGRDRLALNCGSDDEDETEKLTATDONAKITTKKKGGKSLQVLTAILLSQMG
           ASRDALVALVAFRTALSNRLKARVOGLFRAAFRLRAPFLRLSINIMLCFMLS
           TLHSTSKITGALSRLRKILTKIHSVFENNYKISHVDGRITHPEORIASDVR
           FSELSIDLDLTAVTDGILYAWRLCSYAPKYIFMILAYVLGAGTAINFSPSFGK
           LMSKQQLGEYRQLHSRLRTHSEIAFYGGTRESHIOOKFNKLVSHMSHVLDSHW
           WFGMTDGLLKYLGATVAVILIIPEFSGHLRPDDSTLGRAEMISNRVYSHVTSILF
           OALGTLSLRRNLRSYADRHELMVSRRELSDDKSSFORNRSRNLSEANVEF
           SDKVTPTGNVLVEDLTRVEOGSNLLITGGLPYSGHIVKPGVGSOLNKEIYPP
           ORPYMAYGTLRDQLIYPLTSQSESELLTEIGNVELLKKNVLSLFTYRNRVDELLD
           RQPEKRVNWDGSLGEQRLGMRFLYHKPKFALDECTSAVTDDMEERFAAKYRA
           MGTSCITISHRPALVAFHDVVLSDGEGSVHYKRRDSDALLDAEIDSVKSDTDRO
           NDAMVVRFAAARVSAEVLFEFSA7NSKAQSYOTLIARSPVVDKSVLPRFPQOT
           SORALPSRAAMLNVLPPMKPHLLFCISTWLTYSACADANIDITLLITDHTVTFQI
           PTIFDQGAQLLAVCLVSVRSLISDRIASLNGTTVKYVLEQDKAARVRLIGLSVLQS
           GASSITAPASHLRTORLALGWRIRLTORHLRNLNNAFYKVPHMSNSIDADORLTR
           DEKLTADLGLLGMVKPSVDILWVRYVGLISNICPMRGVATLYTYMLLGLGLR
           VAPDGLDAGEEOLEGKFRPTSSCLCPMEHRLNTHAESIAFPGGAREKAVSELIA
           LAIAAGFWVYFESKLPFVNLELFYHLCDDGLFNSNVASLMVDKFRALLDSMLLR
           KHLWGLIDDFVTQLPNNVTWGLSLLYALEHKGRALVSTQVKYLRVLMMLLIPGE
           LAHLRYLASVVSOFMAFGDILELHKFLELSGGINRIFELDEFDASQSVTENQ
           TRLDSDQLSFSEVDIITPAQKMAKSLCISIVSGSKSLVLTGPNSGKSTSVFRVLRD
           IWPVTGGLRTPSLDINKELGSGNGMFVFPQPYTCGLGTDRDQIYPLSVEEAKRAAK
           LVTSGSETEAGSILDSHLKLTILENVRVLYLLERDVGGWDATTNWEDILSLGEQRLG
           MVSLL"
intron    16301..16393
           /gene="AT4g39850"
           /number=1
exon      16394..16679
           /gene="AT4g39850"
           /number=2
intron    16680..16980
           /gene="AT4g39850"
           /number=2
exon      16981..17205
```

```
/gene="AT4g39850"
/number=3
intron    17206..17340
           /gene="AT4g39850"
           /number=3
exon      17341..17538
           /gene="AT4g39850"
           /number=4
intron    17539..17674
           /gene="AT4g39850"
           /number=4
exon      17675..18129
           /gene="AT4g39850"
           /number=5
intron    18130..18221
           /gene="AT4g39850"
           /number=5
exon      18222..18357
           /gene="AT4g39850"
           /number=6

Alignment Scores:
Pred. No.:      2,37e-19      Length:      112067
Score:          629.50      Matches:      145
Percent Similarity: 47.08%      Conservative: 24
Best Local Similarity: 40.39%      Mismatches: 31
Query Match:     34.47%      Indels:      160
DB:              8          Gaps:         5

US-09-720-384a-4 (1-343) x ATCHRIV92 (1-112067)
Qy  142  LysSerThrAsnIleLeuTprHisAsnCysLeuIleGlyGlnSerAspArgGlnLysLeu 161
      |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
Db  47219 AAGCAGAGAACATCGTGTGGCAGAGAGTTCGATATGCAGATCGCAGACACAAACATT 47278

Qy  162  LeuGlyGlnLysGlyCysValValTprIleThrGlyLeuSerGlySerGlyLysLeuThr 181
      |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
Db  47279 CTTCAACAAAGGGTGTGTGCTTGTGATCACTGCTCACTGCTTCAGGAAACACT 47338

Qy  182  LeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTyrrValLeuAsp 201
      |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
Db  47339 GTTGCTGTGCACCTAAAGTAAAGCAATGTTGTTAAAGAGGCAAACTTACTTACACACTCGAC 47398

Qy  202  GlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLysAlaGluAspArgAla 221
      |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
Db  47399 GGCACAATGTCCTCAGCGCCTTACCCTGACCTCACTTTCAAAGCAGAGACCACCAC 47458

Qy  222  GluAsnIleArgArgVal----- 227
      |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
Db  47459 GAAACATTAGAAGAATTTGTACACAAATTTTCTCTTCTTCCAAAAATTTATAATCATGTG 47518

Qy  227  ----- 227
      |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
Db  47519 TTTACTAACTCATGAACCAATTTCTTAAACCGCTTTTAAACCGCTTTTATGCTATATGCTTAATGCTTTT 47578

Qy  228  -----GlyGluValAlaLysLeuPhe 234
      |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
Db  47579 TGTAAAGTGATTTGTGTGATGATCTTTAAATTTAATATATCAGGTGAGTGAAGTTGTTT 47638

Qy  235  AlaAspAlaGlyValIleCysIleAlaSerLeuIleSerProTyrrArgArgAspArgAsp 254
      |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
Db  47639 GCTGAGCTCGGAGTCATTGTATAGCAAGTTTGAATTTCTCGTACCAGCAGACAGACAGAC 47698

Qy  255  AlaCysArgAlaLeuLeuProHisSerAsnPheIleGluVal----- 268
      |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
Db  47699 GCGTCCGGTCTCTTGTATTACCTGACGCGGATTTTCGTCGAGGT-ACGTACGCTATAACGTTT 47757

Qy  268  ----- 268
      |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
Db  47758 TACTATATCAATTTGCTTGTGTTATAGAGATATATTAGTATACATTGATTACATTACGA 47817

Qy  269  -----PheIleAspLeuProLysIleCysGluAlaArgAsp 281
      |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
```

Db 47818 AAACCTTACTTTGATTAGTCTTCATGAGGTTCTCTTCATGTCGCCAGTCGAGAT 47877

QY 282 ProLysGlyLeuTyrLysLeuAlaArgThrGlyLysIleLys----- 295

Db 47978 CCAAGGGGTTGTACAAGCTTGCACGTGCAGGCAAAATCAAAGGTATCCCAAAAAAAGA 47937

QY 295 ----- 295

Db 47938 AAACGTACTTGTATATAAAGTTGACTTTTGTACCTTTTATTTTGTTCGATTCGTAGATTGA 47997

QY 295 ----- 295

Db 47998 CTTATTGACCTAATGTTGAGTGTAACCTTATTTCACCTAGCTAATGTTGAGTTTCTTGA 48057

QY 296 -----GlyPheThrGlyIleAspAspProTyrGluProProIleAsnGlyGlu 311

Db 48058 ATTAATTTACAGGCTTCACTGGAATCGACGACCTTACGAGGACCACTGAATTCGAG 48117

QY 312 Ile-ValIleLysMet----- 316

Db 48118 GTTCGTCACGTGAATGATCAGATAGATTTTGGACGTAATTTGTACTAATAAGTATTG 48177

QY 317 -----LysAspGluGluCysProSerPr 324

Db 48178 ATGTAATATCTCCAGGTAGTCTGAAACACACAGCAGCAGCAGCTGCTGCT---TCGCC 48234

QY 324 oLysAlaMetAlaLysGlnValLeuCysTyrLeuGluGluAsnGlyTyrLeuGln 342

Db 48235 ACCTCAGATGCTGAGAACATCATCTCTTACCTGCACAAACAAAGGTTATCTTCGAG 48289

RESULT 12

ATT5J17

LOCUS ATT5J17 118267 bp DNA linear PLN 21-SEP-1999

DEFINITION Arabidopsis thaliana DNA chromosome 4, BAC clone f5J17 (BSSA project).

ACCESSION AL035708

VERSION AL035708.2 GI:5918309

KEYWORDS

SOURCE thale cress.

ORGANISM Arabidopsis thaliana

REFERENCE 1 Arabidopsis thaliana

2 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids. 1 (bases 1 to 118267)

3 Bevan.M., Murphy.G., Ridley.P., Hudson.S., Bancroft,I., Mewes.H.W., Mayer.K.F.X., Lemcke.K. and Schueller.C.

4 Unpublished

5 EU Arabidopsis sequencing project.

6 Submitted (21-SEP-1999) MIPS, at the Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail: schueller@mips.biochem.mpg.de,mayer@mips.biochem.mpg.de

7 Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK, E-mail: michael.bevan@bbsrc.ac.uk

8 On Sep 22, 1999 this sequence version replaced gi:4490734.

9 Information on performance of analysis and a more detailed annotation of this entry and other sequences of chromosomes 3, 4 and 5 can be viewed at: <http://www.mips.biochem.mpg.de/proj/thal/>.

10 Location/Qualifiers

1. 118267

/organism="Arabidopsis thaliana"

/variety="Columbia"

/db_xref="taxon:3702"

/chromosome="4"

2. 22789

misc_feature 1..22789

/note="position 1-22789 overlaps to BAC t19p19, EMBL accession AL022605; please refer to this entry for analysis and annotation"

22854..24209

/gene="T5J17.10"

22854..24209

exon

misc_feature 26106..26772

/note="EST N95877 matches to positions 26106-26355, 26723-26772; possibly related to subsequent gen"

26806..26867

/gene="T5J17.20"

exon

26806..26867

/number=1

26806..26867

/gene="T5J17.20"

gene

26806..26867

/number=1

26806..26867

/gene="T5J17.20"

CDS

26806..26867

/note="similarity to 70K peroxisomal membrane protein, Rattus norvegicus, PIR2:A35723

Contains ABC transporters family signature [LSLGEQQRLLGMARLF][LSLGEQQRLLGMVSLI]"

/codon_start=1

/product="putative protein"

/protein_id="CA838898.1"

/db_xref="GI:4490736"

/translation="MPSLQLQLTERGRLVASRRKSLIAAGIAGTAVYLSRY ASRRPSSRLNGQSDDETLKLTATDONAKITTKKGGKGLKQLVLTALLSQMG KMGARDLLALVATVFTALSNRLAKVQGLFRAAFLRRAPFLRLISINIMLCFMLS TLHSTSKYITSLRFLKTLKTIHSHYFENMVYKISHWDGRITHPQRLASDVPR FSSLSLLILLDLTAVTDGILYAWRLCSYASPKYIFWILAYVLGAGTAIRNPSFGK LMSKPOOLEGEYRQLHSLRTHSEIAFVGGETREESHIOQKFNLSHMSVLDHWF WFGMTQDELLKYLGAIVAVLLIEPFESCHLRPDSDTLGRAEMLNIRYHTVILSLF QALGTLSISSRLNRLSGADRIHELMVAVSRELSDGKSSQFQRNRLNLSANKVEF SDVKVPTGTVNLVEDTLRVEQGNLLITGLMPLVSGHIVKPGVSDLNKEIFYVP QRYMAVTLRDLDIYPLTSGOSELLEIEMGVELLNKNSILVSYFTRMVDEYLLD RYQPEKENVNGDELSLGEQQRLLGMARLFYHKPFALIDECTSAVTDMEEFAAKVRA MGTCTITSHRPALVAFHDVLSLDGEGSVHYKRDSALLTDABIDSVKSDTDRO NDAMVORFAAARKVSAEVLFEASATNSKAOSYOTLIARISPVVDSVLPFPQPOF SORALPSRVAAMLNVLNPKPHLLFCISTLTYSACADANIDILLKTIHDFIVFOI PTIPDKOALLVACLVVSRTLISDRASINAGTVKLVLEODKAAFLVGLSLVLOS GASSIAPSRHLHTQRLALQWRILRTOHLNRLNNAFYKVFHMSGNSIDADQRLTR DLEKLTADLSGLLTGMVKSVDILMTYRVGLISNICTMRGVAILTYMLLGLFLRR VAPDGLDEGEQQLKEPYNLELFYHLCDDGLFNSVNASLMVDKRFKRALLDHSLMLR LAIAGFWVWIFESKLPENLELFYHLCDDGLFNSVNASLMVDKRFKRALLDHSLMLR KKWLYGLDDFTVKLPNNVTWGLSLLYALEHKGDRALVSTQVKYLVHLYMLLIPGE LAHALRYLASVVSQSFMAFGDILELHKFLFELSGGINRIFELFELDASQSGVTSENO TSRLSDODLLSFSEVDITPAOKLMASKLSCEIVSGSKLLVTGTPNGSGKTSVFRVLR TWPTVCGRLTKPSLIDIKELGSGNGMFFVQRPYTCGLTLDQOILYPLSEAEKPAK LYTSGESSTAGSILDSHLKTLILENRVLYLLERDVGWDATTNWDITLSLGEQQRUG MVSLL"

26868..26960

/gene="T5J17.20"

/number=1

26961..27246

/gene="T5J17.20"

/number=2

27247..27547

intron

intron

Db 58505 AAACGTACTGTATATAAAGTTGACTTTGACTTTTATTTTGTTCGATTCGTAGAGTTTCA 58564
QY 295 -----
Db 58565 CTTATTGACCTAATTGTTGAGTGTAACTTATTGACCTAGCTAAATTGTTGAGTTTCTTCA 58624
QY 296 -----GlyPheThrGlyLeuAspAspProTyrGluProProlleAsnGlyGlu 311
Db 58625 ATTAATATTCAGGCTTCTACCTGGAATCGACGACCTTACGAGGACCAAGTGAATTCGAG 58684
QY 312 ile-ValileLysMet-----
Db 58685 GTTCGTACAGTCAATGATCAGATAGATTGGAACGCTAAATTGTTACTAATAAAGTATTG 58744
QY 317 -----LysAspGluGluCysProSerPr 324
Db 58745 ATGTAATATCCTCCAGGTAGTGTGGAACACACAGGACGAGGTCGTGT---TCGCC 58801
QY 324 oLysAlaMetAlaLysGlnValLeuCysTyrLeuGluGluAsnGlyTyrLeuGln 342
Db 58802 ACCTCAGATGGCTGGAACATCATCTCTTACCTGCACAAACAAAGGTTATCTTGAG 58856
RESULT 13
AF069952 1017 bp mRNA linear PLN 20-JUL-1998
LOCUS Enteromorpha intestinalis 5'-adenylylsulfate kinase mRNA, complete cds.
DEFINITION
ACCESSION AF069952
VERSION AF069952.1 GI:3329470
KEYWORDS hollow green seaweed.
SOURCE Eukaryota; Viridiplantae; Chlorophyta; Ulvophyceae; Ulvales;
ORGANISM Ulvaceae; Enteromorpha.
REFERENCE
AUTHORS Gao,Y. and Leustek,T.
TITLE Cloning of the cDNA for 5'-adenylylsulfate (APS) kinase from the marine macroalga Enteromorpha intestinalis
JOURNAL Unpublished
AUTHORS Gao,Y. and Leustek,T.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1998) Biotech, Rutgers University, 59 Dudley Road, New Brunswick, NJ 08901-8520, USA
FEATURES
source 1..1017
/organism="Enteromorpha intestinalis"
/db_xref="taxon:3116"
/EC_number="2.7.1.25"
/function="catalyzes the phosphorylation of APS to 3'-phosphoadenosine, 5'-phosphosulfate (PAPS)"
/note="EAPK1"
/codon_start=1
/product="5'-adenylylsulfate kinase"
/protein_id="AAC26856.1"
/db_xref="GI:3329471"
/translation="MLRAIAQRARGSLALQCAPTEWASCVRGSGGTAYDVGEGSTNI KWHETVSRGDKERLLNORGVLFWTGLSGSGKSTVACTLEHALNARGKMTALLDGN KWHETVSRGDKERLLNORGVLFWTGLSGSGKSTVACTLEHALNARGKMTALLDGN EVECIMKPIELCEQRDPKGLYKARGLMKMGFTGIDDPYEEFPELFTVYREGGSD MNSPEAKAKOIFDYLEAKGLKGPVAVSGSGCARVARGHGRRRGRORLAW"
BASE COUNT 193 a 313 c 353 g 158 t
ORIGIN
Alignment Scores:
Pred. No.: 7, 95e-20 Length: 1017
Score: 595.50 Matches: 126
Percent Similarity: 69.60% Conservative: 45
Best Local Similarity: 51.43% Mismatches: 62
Query Watch: 32.61% Indels: 13
DB: 8 Gaps: 4

US-09-720-384A-4 (1-343) x AF069952 (1-1017)
QY 108 GlyIleGlyArgTrpValArgArgArgArgAsnGly-----AlaAlaPro 123
Db 22 GGGATGCTGGAGCATCGCAACGGCGCGCTCGCGCGCTCAATGCGCGCGCG 81
QY 124 GlyGluAlaProHisSerProValLysGluLysProValMetSer-----AsnIleGly 141
Db 82 GGCACCGAGTGGCCGACGTCGCTACGCGGATCGTCGGGCTTCACGCTATGATGTCGGC 141
QY 142 LysSerThrAsnIleLeuThrHisAsnCysLeuIleGlyGlnSerAspArgGlnLysLeu 161
Db 142 GAGTCGACCAACATCAAGTGGCAGCAGATGTCGAGCGCGGCGGCGACAAAGGAGCGCTG 201
QY 162 LeuGlyGlnLysGlyCysValValTrpIleHcIleLeuSerGlySerGlyLysSerThr 181
Db 202 CTCAATCAGGCTGGCTGCGCTCTCTGCTTCACAGGCTCTCTGCTTCGCGCAAGTCCACT 261
QY 182 LeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTyrValLeuAsp 201
Db 262 GTGGCTGTGCACGCTCGAGCACGCCCTGAACGCGCGCGCAAGATGACGCGCGCTGCGAC 321
QY 202 GlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLysAlaGluAspArgAla 221
Db 322 GCGCACCAACGTTCCGACCGCTTGAACCTCAAACTTCACACTGAGGACAGAA-- 379
QY 222 GluAsnIleArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyValIleCys 241
Db 380 CGGAACATCGCGCGCATCGCGGAAGTGAAGCTCTTTCGCGACGCTGCGCTCCAC 439
QY 242 IleAlaSerLeuIleSer-ProTyrArgArgAspArgAlaCysArgAlaLeuLeuPr 261
Db 440 TCCGTGACCTTCATCTCCGCCCTATCGCGCGCGCACCCCGCGCGAGCGCTG---- 494
QY 261 oHis-SerAsnPheIleGluValPheIleAspLeuProLeuLysIleCysGluAlaArgA 281
Db 495 -CGCGCGCGACTTCGTGGAGTCTATATGAAGATCCCCATCGAGCTCTCGAGCAGCGCG 553
QY 281 spProLysGlyLeuTyrLysLeuAlaArgThrGlyLysIleLysGlyPheThrGlyIleA 301
Db 554 ACCCAAGGGCCCTTACAAGAAGCGCGCGCGCTCATGATGAGGCTTCACCGCATCG 613
QY 301 spAspProTyrGluProProlleAsnGlyGluIleValIleLysMetLysAspGlu---- 319
Db 614 ACGACCATACGAGAGCGCGCTGGAGCTGAGCTACGATCAGGTCGCGGAGGAGGCA 673
QY 320 --GluCysProSerProLysAlaMetAlaLysGlnValLeuCystyrLeuGluGluAsnG 339
Db 674 GCGACATGAACCTCGCCGAGGCGCATGCGAAGCAGATCTTTGACTATCTCGAGGCGAAG 733
QY 339 lyTyrLeuGln 342
Db 734 GCTTCTCAAG 744
RESULT 14
ATAC009540 101410 bp DNA linear PLN 24-JAN-2001
LOCUS Arabidopsis thaliana chromosome III BAC F20H23 genomic sequence, complete sequence.
DEFINITION
ACCESSION ATAC009540
VERSION ATAC009540.6 GI:12408711
KEYWORDS HTG.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE 1 (bases 1 to 101410)
AUTHORS Lin,X., Kaul,S., Town,C.D., Benito,M.-I., Creasy,T.H., Haas,B., Ronning,C.M., Koo,H., Fujii,C.Y., Utterback,T.R., Barnstead,M.E., Bowman,C.L., White,O., Nierman,W.C. and Fraser,C.M.
TITLE Arabidopsis thaliana chromosome III BAC F20H23 genomic sequence


```
/protein_id="AAF00627.1"
/db_xref="GI:6006851"
/translation="MNAAATNRNFKLASRLDLGLSKLEKSLIPFRELIVECTIPKD
DGTIASFGRVQNDNAGPKMGIRHPEVEDEVNALAQMLNWKTAVALKIPYGAK
GGICDSELSLELUTRVFTQKIHDLIGIHDPADMDGTGPQMAWILDEISKF
HGSPAVVYTKPDLGGSLGRDAATGRVLFAATEALLNEHGKTI SGOQFAIQGFQNGV
SWAAKISDKGKIVAVSDVTGAIKNNNGDIILSLEHAENRGIKGFDSADSIDPS
ILVDDCDILVPALAGVINRENANEIRAKFIIEGANHPDTPDEADILKKGVMILPDI
YANGSGTVSYFEWVIQGFQFMDDEEKVYMRGFKDLKEMCQTHSCDLRMDA
FTLGINVAQATTIRCGWS"
10160..10189
/note="exon predicted by xgrail, quality good_shadowexon"
11181..111304
/note="exon predicted by xgrail, quality
marginal_shadowexon"
11576..11597
/rpt_family="POLY_A"
complement(12077..12132)
/note="exon predicted by xgrail, quality
excellent_shadowexon"
12685..12742
/note="exon predicted by xgrail, quality marginal"
join(<12760..12883,12994..13143,13397..13515,13609..13693,
13804..13850,13955..>14056)
/gene="F20H23.5"
<12760..>14056
/gene="F20H23.5"
/note="similar to GB:S47640 [Arabidopsis thaliana]"
join(12760..12883,12994..13143,13397..13515,13609..13693,
13804..13850,13955..14056)
/gene="F20H23.5"
/codon_start=1
/product="putative adenylylsulfate kinase"
/protein_id="AAF00628.1"
/db_xref="GI:6006852"
/translation="MSTVGNSTNIFWQESPITGRQKLLNKGCVWITGLSGSKS
TLACSLRELNRRGLSYILDGNLRHLNKLGLFKAEDVENIRRVGEVAKLPADAG
LICTASLISPYRKDRDREMIONSSFTEVFMNMSQLCEARDPKGLYKLARAKIKG
```

Alignment Scores:

```
Pred No.: 1 86e-17 Length: 101410
Score: 587.50 Matches: 151
Percent Similarity: 41.76% Conservative: 29
Best Local Similarity: 35.03% Mismatches: 26
Query Match: 32.17% Indels: 228
DB: 8 Gaps: 5
```

US-09-720-384A-4 (1-343) x ATAC009540 (1-101410)

```
QY 137 MetSerAsnIleGlyLysSerThrAsnIleLeuTrpHisAsnCysLeuIleGlyGlnSer 156
||||| :||| ||||| ||||| ||| ||||| :|||
DB 12760 ATGTCGACAGTGGGAATTCACGACACATATTTTGGCAAGAAATCCCCATTTGGGAAACT 12819

QY 157 AspArgGlnLysLeuGlyGlnLysGlyCysValIleThrGlyLeuSerGly 176
:::||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 12820 GAAAGGCAGAGAGTTGTTAAATCAGAGAGGTTGTGTGGATCACAGGCTCAGTGGC 12879

QY 177 Ser----- 177

DB 12880 TC-AGGTCAATCTTCTAAATAATCTTAAATCTTTTCCCATGTGATACATTACTAAGAT 12938

QY 178 -----GlyLys 179

DB 12939 GCCTTCTTTATCTCTGCTATATATCTTAAATCTTTAGCCTCAAACTTTGTGCAGGCAAA 12998

QY 180 SerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTyrVal 199
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 12999 AGCACGTAGCTTGTCTGCTTACTAGAGAGCTGAACAACGGGGAGAACTATCATATATT 13058

QY 200 LeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLysAlaGluAsp 219
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 13059 CTTGATGTCACAAATCTTCGTCATGGTTTGACAAAGAGATCTTGGTTTCAAGGCAGAGAT 13118
```

```
QY 220 ArgAlaGluAsnIleArgArgVal----- 227
||| ||||| ||||| ||||| |||||
DB 13119 AGAGTCGAAAATATACCGCAGGTCGTAAGTAGTACATAGTTGCTTACTATCTTTTGCCT 13178

QY 227 ----- 227

DB 13179 TCAGTTGATGAAAGTCTGATTCCGAGAAACTCAGTAATACAAATGCAAAATCCTTTAAAA 13238

QY 227 ----- 227

DB 13239 AGGTGCTTCATATGATAGTTTACATGGTACCAATACCTCATTTGATATACATATTG 13298

QY 227 ----- 227

DB 13299 GTATACACCTAAAGCTGATAGATTCTGGTTAATGTTGAAAGAACTTCTCTAAAGGCCA 13358

QY 228 -----GlyGluValAlaLysLeuPheAl 235
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 13359 AAATACTAAAGAGAACCAACAATTTGTTCTTAAATGCAGGAGAGTAGCCAAACTCTTTGC 13418

QY 235 aAspAlaGlyValIleCysIleAlaSerLeuIleSerProTyrArgArgAspArgAspAl 255
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 13419 GGATGCTGTTTAATCTGATTGCCAGCTCATATCCCGTATAGAAAAAGACCGTGACGC 13478

QY 255 aCysArgAlaLeuLeuProHisSerAsnPheIleGlu----- 267
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 13479 CTCAGGGAATGATACAGAATTCATCTTTATTGAGCAAGCTATTCTAAACACCGCTT 13538

QY 267 ----- 267

DB 13539 TACAGTTGTAACCCCATTCGGTTGGATCTTAAAAACATCTTTGTTGATATTGTGGCCA 13598

QY 268 -----ValPheIleAspLeuProLeuLysIleCysGluAlaArgAspProLysG 284
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 13599 AATTGTTGAGTTTTCATGACATGCTCTGCAATTTGTGTGAAGCAGGACCCCTAAAGG 13658

QY 284 yLeuTyrLysLeuAlaArgThrGlyLysIleLys----- 295
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 13659 CCTATACAGCTTGCACGTGCAGGAAAAGATCAA-AGGTGAGATAACATAGCTTCAAGGTT 13717

QY 295 ----- 295

DB 13718 CCATATAACAGACAGATAGCAAGTTGCTTCAATATCATTTGAGTTGTTCTATATAAGTTT 13777

QY 296 -----GlyPheThrGlyIleAspAspProTyrGluProPr 307
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 13778 TTTTTCCTTCATACCTTTTGTGTAGGTTTCACAGGAATAGATGATCGGTATGAGTCTC 13837

QY 307 oIleAsnGlyGluIleVal----- 313
|:|:|:| |||||
DB 13838 CTTGAACCTGTGAGGTATG-TGATCCATCTGAACATAAGATTCTTCTCAAGAATATTGCA 13896

QY 313 ----- 313

DB 13897 AACAAACCGGTATGTAATACGTTCTCTGTTCTTTTGTATGTTTTTGAACATAACAGAT 13956

QY 314 ----IleLysMetLysAspGluGluCysProSerProLysAlaMetAlaLysGlnValLe 332
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 13957 AGAGTTCAAAGAGAAAGAGGAGAGTGTCCTTCCCTGTAGCTATGCTGAGGAAGTGAT 14016

QY 332 uCysTyrLeuGluLysAsnGlyTyrLeuGln 342
: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 14017 CTCCTTATCAGAGACAAAGGTTTCCCTTCAA 14047
```

RESULT 15

```
LOCUS AE007523 11131 bp DNA linear BCT 27-JUL-2001
DEFINITION Clostridium acetobutylicum ATCC824 section 11 of 356 of the
complete genome.
ACCESSION AE007523 AE001437
VERSION AE007523.1 GI:15022925
KEYWORDS
SOURCE Clostridium acetobutylicum.
```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 2, 2002, 01:54:13 ; Search time 197 Seconds
(without alignments)
2989.348 Million cell updates/sec

Title: US-09-720-384A-4
Perfect score: 1826
Sequence: 1 RFFHFINQTEPLVHTTQQPP.....PKAMAKQVLCVLENGYLQA 343

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1736436 seqs, 858457221 residues
Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q/cgn2.1/USPTO.spool/US09720384/runat_30102002.090527.19173/app_query.fasta_1.519
-DB=N_Geneseq_032802 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cd1
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09720384.cgn.1.1.145 @runat_30102002.090527.19173 -NCPU=3
-NO_XLPXY -NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_032802.*

1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001C.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1826	100.0	1217	21	AAZ50160
2	887	48.6	890	21	AAZ50159
3	876	48.0	928	21	AAZ50163
4	852	46.7	936	21	AAZ50162
5	834	45.7	627	21	AAAC44691
6	817	44.7	1175	21	AAC36588
7	816	44.7	1068	21	AAC35266
8	816	44.7	1183	21	AAC45659
9	809.5	44.3	714	21	AAZ51190
10	778.5	42.6	917	21	AAC98337
11	778.5	42.6	948	21	AAC42294
12	757.5	41.5	1130	21	AAC49070
13	752.5	41.2	1133	21	AAC36627
14	617.5	33.8	483	21	AAC51315
15	502.5	27.5	2646	24	AAS94855
16	496	27.2	2507	19	AAV33482
17	495	27.1	1875	20	AAV84897
18	495	27.1	2527	20	AAK00717
19	495	27.1	2527	20	AAK00686
20	494.5	27.1	431	21	AAZ50161
21	489.5	26.8	990	16	AAT04849
22	486.5	26.6	2617	21	AAZ94206
23	484.5	26.5	2990	22	AAS33174
24	484	26.5	650	21	AAF14521
25	481	26.3	606	22	AAH81465
26	475	26.0	1160	21	AAA08597
27	454.5	24.9	868	21	AAF07688
28	419	22.9	588	21	AAF1283
29	390	21.4	582	21	AAZ49624
30	333	18.2	347	22	AAH42882
31	314.5	17.2	34980	22	AAH41225
32	297.5	16.3	3332	23	AS88909
33	255.5	14.0	447	21	AAC35896
34	249	13.6	521	21	AAZ50164
35	202.5	11.1	195	22	ABA71730
36	202.5	11.1	195	22	ABA37826
37	202.5	11.1	195	22	AAK20089
38	202.5	11.1	195	22	AAK46143
39	202.5	11.1	195	22	AAI52052
40	189	10.4	471	22	ABA59209
41	189	10.4	471	22	ABA27965
42	189	10.4	471	22	AAK07411
43	189	10.4	471	22	AAK33189
44	189	10.4	471	22	AAI38986
45	184	10.1	53500	21	AAC55842

ALIGNMENTS

RESULT 1
AAZ50160
ID AAZ50160 standard; cDNA: 1217 BP.
XX
AC AAZ50160;
XX
DT 04-MAY-2000 (first entry)
XX
DE Corn Adenylsulphate kinase-2 cDNA clone.
XX
KW Adenylsulphate kinase; Adenosine-5'-phosphosulphate kinase; APS kinase;
KW 3'-Phospho-adenosine-5'-phosphosulphate; PAPS; sulphate assimilation;
KW corn; clone p0016.ctscj40rb; transgenic plant; screen; antibody; ss.
OS Zea mays.
XX
FH Key
FT CDS 2..1033
Location/Qualifiers

FT /**tag= a
 FT /product= "Corn APS kinase-2"
 FT /note= "Derived from clone p0016.ctscj40rb"
 XX
 PN W0200004165-A1.
 XX
 XX 27-JAN-2000.
 XX
 XX 13-JUL-1999; 99WO-US15809.
 XX
 XX 14-JUL-1998; 98US-0092833.
 XX
 XX (DUFO) DU PONT DE NEMOURS & CO E I.
 XX
 XX Falco SC, Allen SM, Anderson SL;
 XX
 XX WPI; 2000-182430/16.
 DR P-PSDB; AAY44789.
 XX
 XX New nucleic acid molecule and chimeric gene encoding an adenosine-5'
 PT phosphosulphate kinase, useful for altering expression of sulfate
 PT assimilation protein in plants -
 XX
 XX Claim 3; Page 30; 42pp; English.
 XX
 CC The present sequence is a cDNA encoding corn adenylsulphate kinase
 CC (APS kinase), also known as adenosine-5' phosphosulphate kinase. This is
 CC obtained from clone p0016.ctscj40rb, derived from corn pooled tassels
 CC shoots, p0016 cDNA library. APS kinase is a sulphate assimilation
 CC protein, that catalyses the conversion of adenosine-5' phosphosulphate
 CC (APS) to 3'-phospho-adenosine-5' phosphosulphate (PAPS). This sequence
 CC is used as probe and primers to identify, obtain and synthesise sulphate
 CC assimilation proteins from other plants. It is also used to produce
 CC transgenic plants, that are useful for altering the expression levels of
 CC a sulphate assimilation protein. The APS kinase peptides are useful for
 CC producing antibodies, that are used to screen and isolate cDNA clones.
 XX
 SQ Sequence 1217 BP; 344 A; 292 C; 303 G; 278 T; 0 other;

 Alignment Scores:
 Pred. No.: 1.32e-104 Length: 1217
 Score: 1826.00 Matches: 343
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 21 Gaps: 0

 US-09-720-384A-4 (1-343) x AA250160 (1-1217)

 QY 1 ArgProPheHisPheIleAsnGlnThrGluProLeuValThrHisThrGlnProPro 20
 Db 2 CGTCCGTTTCATTTCATCAATCAACAGAACCTCTGGTCACACACAGCAGCAGCCG 61

 QY 21 SerProAlaProGlyProAlaSerGlnGlyGlnArgGlnGlyAsnThrLeuLeuSerPro 40
 Db 62 AGCCAGCGCCGCCGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCC 121

 QY 41 ThrProThrLeuAlaValIleLeuValAsnProGlnArgAlaProProValLeuProGly 60
 Db 122 AGCCGACGCTCGCGGTCATCTCGTAATCCACAGCGCGGCTCCGCTCTCCAGGC 181

 QY 61 LeuThrProSerAspAlaProLeuProAlaLeuValIleHisGlyLeuThrProArgSer 80
 Db 182 CTCACCCCTAGCGATGGCCACTCCCGCGGCTCGTGATCCATGGCTCCTACTCCCGTCC 241

 QY 81 SerHisSerSerAlaGlyLeuAlaSerAspSerGlyArgGlyGlyGlyArgGly 100
 Db 242 TCACACTCTTCGCGGGTCTCGCCAGTGTAGTGGGCGCGCGAGGGGAGGGCGCGGT 301

 QY 101 AlaArgThrHisCysHisArgGlyIleGlyArgTrpValArgArgArgArgAsnGly 120
 Db 302 GCGCGTACGCACTGCCACCGCGGCATTGGCGGTTGGGCGGCGCGCGCGCGAATGGA 361

QY 121 AlaAlaProGlyGluAlaProHisSerProValLysGluLysProValMetSerAsnIle 140
 Db 362 GCAGGCGCGCGGGAGCGCGCCAGCAGCAGGAGAGAGCTGTAAATGTCGAACATT 421

 QY 141 GlyLysSerThrAsnIleLeuTrpHisAsnCysLeuIleGlyGlnSerAspArgGlnLys 160
 Db 422 GGGAAATCGACTAATATTTATGGCAACAATTCCTTGATGGACAATCTGATACAGAAA 481

 QY 161 LeuLeuGlyGlnLysGlyCysValValTrpIleThrGlyLeuSerGlyLysSer 180
 Db 482 TTGCTGGGACAAAAGGCTGTGCTGATGATACAGGACTCAGTGGTTCAGGGAAAGT 541

 QY 181 ThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTyrValLeu 200
 Db 542 ACTCTTGCTGTCGACTGAGTGTGCTGATTCGAGAGCCACCTCAGTATGTACTT 601

 QY 201 AspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLysAlaGluAspArg 220
 Db 602 GATGCTGACACCTCAGACATGGCCTAAATAGAGATTTAAGCTTTAAGGCAGAACCGT 661

 QY 221 AlaGluAsnIleArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyValIle 240
 Db 662 GCAGAAATATACGAGAGTGGTGAAGTGGCAAGCTTTTGTCTGATGCTGTGTCATA 721

 QY 241 CysIleAlaSerLeuIleSerProTyrArgArgAspArgAspAlaCysArgAlaLeuLeu 260
 Db 722 TGCATTTGCTAGCTTTGATATCTCCATACAGAGAGATCGTGATGCGCTGCTACTT 781

 QY 261 ProHisSerAsnPheIleGluValPheIleAspLeuProLeuLysIleCysGluAlaArg 280
 Db 782 CCACATTTCACTTATTTAAGTATTTATGATTTGCCCTTAAAAATTTGTGACGCTGT 841

 QY 281 AspProLysGlyLeuTyrLysLeuAlaArgThrGlyLysIleLysGlyPheThrGlyIle 300
 Db 842 GATCCTAAAGCGCTATACAAAGCTTGACGTCAGGAGAAAGATTAAGGTTTCACCTGGAAT 901

 QY 301 AspAspProTyrGluProProIleAsnGlyGluIleValIleLysMetLysAspGluGlu 320
 Db 902 GATGATCCATACGACCAACCAATTAATGCTGAGATAGTAATTAAGATGAAGATGAGAA 961

 QY 321 CysProSerProLysAlaMetAlaLysGlnValLeuCysTyrLeuGluGluAsnGlyTyr 340
 Db 962 TGCCCTTCACCAAGCAATGGCCAAAGTTCCTATGCTACCTTGAAGAAAACGGATAT 1021

 QY 341 LeuGlnAla 343
 Db 1022 TTGCAAGCT 1030

 RESULT 2
 AA250159
 ID AA250159 standard; cDNA; 890 BP.
 XX
 AC AA250159;
 XX
 DT 04-MAY-2000 (first entry)
 XX
 DE Corn Adenylsulphate kinase-1 cDNA clone.
 XX
 KW Adenylsulphate kinase; Adenosine-5'-phosphosulphate kinase; APS kinase;
 KW 3'-Phospho-adenosine-5'-phosphosulphate; PAPS; sulphate assimilation;
 KW corn; clone cen3n.pk0088.b10; transgenic plant; screen; antibody; ss.
 XX
 OS Zea mays.
 XX
 FH Key Location/Qualifiers
 CDS 3..677
 FT /**tag= a
 FT /product= "Corn APS kinase-1"
 FT /note= "Derived from clone cen3n.pk0088.b10"
 XX
 PN W0200004165-A1.
 XX
 XX 27-JAN-2000.

```
XX 13-JUL-1999; 99WO-US15809.
XX 14-JUL-1998; 98US-0092833.
XX (DUPO ) DU PONT DE NEMOURS & CO E I.
XX Falco SC, Allen SM, Anderson SL;
XX WPI; 2000-182430/16.
XX P-PSDB; AAY44788.
XX New nucleic acid molecule and chimeric gene encoding an adenosine-5'
XX phosphosulphate kinase, useful for altering expression of sulfate
XX assimilation protein in plants
XX Claim 3; Page 29; 42pp; English.
XX The present sequence is a cDNA encoding corn adenylsulphate kinase
XX (APS kinase), also known as adenosine-5' phosphosulphate kinase. This is
XX obtained from clone cen3n.pk0088.b10, derived from corn endosperm, cen3n
XX cDNA library. APS kinase is a sulphate assimilation protein, that
XX catalyses the conversion of adenosine-5' phosphosulphate (APS) to
XX 3'-Phospho-adenosine-5' phosphosulphate (PAPS). This sequence is used as
XX probes and primers to identify, obtain and synthesise sulphate
XX assimilation proteins from other plants. It is also used to produce
XX transgenic plants, that are useful for altering the expression levels of
XX a sulphate assimilation protein. The APS kinase peptides are useful for
XX producing antibodies, that are used to screen and isolate cDNA clones.
XX Sequence 890 BP; 232 A; 239 C; 262 G; 157 T; 0 other;
SQ
Alignment Scores:
Pred. No.: 1,23e-46 Length: 890
Score: 887.00 Matches: 171
Percent Similarity: 84.23% Conservative: 16
Best Local Similarity: 77.03% Mismatches: 35
Query Match: 48.58% Indels: 0
DB: 21 Gaps: 0
US-09-720-384A-4 (1-343) x AAZ50159 (1-890)
QY 121 AlaAlaProGlyGluAlaProHisSerProValysGluLysGluLysProValMetSerAsnIle 140
Db 6 GCGCGCGCGCGCTGCGAGGATCATGACAGCAGCAGCGCGTGTGTGACCTCGACCGCTC 65
QY 141 GlyLysSerThrAsnIleLeuTrpHisAsnCysLeuIleGlyGlnSerAspArgGlnLys 160
Db 66 GGGAAATCGAGCAACATCTGTGGCATGAGTCCGCCATCGGCGAGAAGGACGACAGGCT 125
QY 161 LeuLeuGlyGlnLysGlyCysValValTrpIleThrGlyLeuSerGlySerGlyLysSer 180
Db 126 CTGCTGAACACAGAGGCTGCGTGTGTGTGATCACTGGCTTAAGCGGTTTCAGGGAAGC 185
QY 181 ThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTrpValLeu 200
Db 186 AGGCTCGCGTGCAGCGTGCAGCGAGCTGCACGCGCAGCCACCTCAGTACGTCCTC 245
QY 201 AspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLysAlaGluAspArg 220
Db 246 GACGCGACACCTCAGGACCGGCTGACAGGACCTCAGCTCGGACGAGGACCGC 305
QY 221 AlaGluAsnIleArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyValIle 240
Db 306 GCGGAGAACATCCGACAGTAGGGGAAGTAGCAAGTGTTCGCCGACGCTGGCTCCTC 365
QY 241 CysIleAlaSerLeuIleSerProTrpArgArgAspArgAspAlaCysArgAlaLeuLeu 260
Db 366 TCATTCGCGACCTCATATGCCCTACAGAGCGGACCGGATCGCGATCTGCTG 425
QY 261 ProHisSerAnPheIleGluValPheIleAspLeuProLeuLysIleCysGluAlaArg 280
Db 426 CCCAAGCACTCGTGTATCGAGGTGTCTCTGGACCTGCGCGCTTCAAGTGTGCGAAGCCAGG 485
```

```
QY 281 AspProLysGlyLeuTrpLysLeuAlaArgThrGlyLysIleLysGlyPheThrGlyIle 300
Db 486 GACCCCAAGGCGCTCTACAAGCTCGCAGCGCGCAAAATCAAGGTTTCACCGGCATC 545
QY 301 AspAspProTrpGluProPheIleAsnGlyGluIleValIleLysMetLysAspGluGlu 320
Db 546 GAGATCCTTACGAACCGCGCTCGGACTGTGAGATAGTATCATCTGTAAGTCGGCAG 605
QY 321 CysProSerProLysAlaMetalLysGlnValLeuCysTrpLeuGluGluAsnGlyTrp 340
Db 606 TGCCTTCGCTCGAATCGATGCTGCTGACGTGTGCTGCTGACCTTGAGACCAATGGTTTC 665
QY 341 LeuGln 342
Db 666 CTCAG 671
RESULT 3
AAZ50163
ID AAZ50163 standard; cDNA; 928 BP.
XX AAZ50163;
AC AAZ50163;
DT 04-MAY-2000 (first entry)
XX Wheat Adenylsulphate kinase-1 cDNA clone.
DE Adenylsulphate kinase; Adenosine-5'phosphosulphate kinase; APS kinase;
XX 3'-Phospho-adenosine-5'phosphosulphate; PAPS; sulphate assimilation;
KW wheat; clone wrl.pk0101.e2; transgenic plant; screen; antibody; ss.
XX Triticum aestivum.
OS
FH Key Location/Qualifiers
FT CDS 3..743
FT /tag= a
FT /product= "Wheat APS kinase-1"
FT /note= "Derived from clone wrl.pk0101.e2"
XX
XX WO200004165-A1.
XX
XX 27-JAN-2000.
XX
XX 13-JUL-1999; 99WO-US15809.
XX
XX 14-JUL-1998; 98US-0092833.
XX
XX (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
XX Falco SC, Allen SM, Anderson SL;
XX
XX WPI; 2000-182430/16.
XX
XX P-PSDB; AAY44792.
XX
XX New nucleic acid molecule and chimeric gene encoding an adenosine-5'
XX phosphosulphate kinase, useful for altering expression of sulfate
XX assimilation protein in plants
XX Claim 3; Page 34-35; 42pp; English.
XX The present sequence is a cDNA encoding wheat adenylsulphate kinase
XX (APS kinase), also known as adenosine-5' phosphosulphate kinase. This is
XX obtained from clone wrl.pk0101.e2, derived from 7 day old light grown
XX wheat root seedlings, wrl cDNA library. APS kinase is a sulphate
XX assimilation protein, that catalyses the conversion of adenosine-5'
XX phosphosulphate (APS) to 3'-Phospho-adenosine-5' phosphosulphate (PAPS).
XX This sequence is used as a probe and primer to identify, obtain and
XX synthesise sulphate assimilation proteins from other plants. It is also
XX used to produce transgenic plants, that are useful for altering the
XX expression levels of a sulphate assimilation protein. The APS kinase
XX peptides are useful for producing antibodies, that are used to screen
XX and isolate cDNA clones.
XX
```


PR	19-JUL-1999;	99US-0144335.
PR	20-JUL-1999;	99US-0144352.
PR	20-JUL-1999;	99US-0144632.
PR	20-JUL-1999;	99US-0144884.
PR	21-JUL-1999;	99US-0144814.
PR	21-JUL-1999;	99US-0145086.
PR	21-JUL-1999;	99US-0145088.
PR	22-JUL-1999;	99US-0145085.
PR	22-JUL-1999;	99US-0145087.
PR	22-JUL-1999;	99US-0145089.
PR	22-JUL-1999;	99US-0145192.
PR	23-JUL-1999;	99US-0145145.
PR	23-JUL-1999;	99US-0145218.
PR	23-JUL-1999;	99US-0145224.
PR	26-JUL-1999;	99US-0145276.
PR	27-JUL-1999;	99US-0145313.
PR	27-JUL-1999;	99US-0145318.
PR	27-JUL-1999;	99US-0145319.
PR	28-JUL-1999;	99US-0145351.
PR	02-AUG-1999;	99US-0146386.
PR	02-AUG-1999;	99US-0146388.
PR	02-AUG-1999;	99US-0146389.
PR	03-AUG-1999;	99US-0147038.
PR	04-AUG-1999;	99US-0147204.
PR	04-AUG-1999;	99US-0147302.
PR	05-AUG-1999;	99US-0147192.
PR	05-AUG-1999;	99US-0147260.
PR	06-AUG-1999;	99US-0147303.
PR	06-AUG-1999;	99US-0147416.
PR	09-AUG-1999;	99US-0147493.
PR	09-AUG-1999;	99US-0147935.
PR	10-AUG-1999;	99US-0148171.
PR	11-AUG-1999;	99US-0148319.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148565.
PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149829.
PR	23-AUG-1999;	99US-0149802.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	28-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161992.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.
Alignment Scores:		
Pred. No.:	1,61e-43	Length: 627
Score:	834.00	Matches: 154
Percent Similarity:	88.35%	Conservative: 28
Best Local Similarity:	74.76%	Mismatches: 24
Query Match:	45.67%	Indels: 0
DB:	21	Gaps: 0
US-09-720-3		

Db 541 AAGAGGGAGAGTGTCTCCCTGTGCTATGCTGAGGAGTGATCTCTTATCTAGAA 600
Qy 337 GluAsnGlyTyrLeuGln 342
::: |||:::|||||
Db 601 GACAAAGGTTTCCTTCAA 618
RESULT 6
AAC36588
ID AAC36588 standard; DNA; 1175 BP.
XX AC AAC36588;
XX XX
DT 17-OCT-2000 (first entry)
XX XX
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 14353.
XX XX
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130410.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 07-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 25-MAY-1999; 99US-0135629.
PR 27-MAY-1999; 99US-0136021.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148563.

XX	06-SEP-2000.	PR	02-JUL-1999;	99US-0142055.
PD		PR	06-JUL-1999;	99US-0142390.
XX		PR	08-JUL-1999;	99US-0142803.
PF		PR	09-JUL-1999;	99US-0142920.
XX	25-FEB-2000; 2000EP-0301439.	PR	12-JUL-1999;	99US-0142977.
PF		PR	13-JUL-1999;	99US-0143542.
PR	25-FEB-1999;	PR	14-JUL-1999;	99US-0143624.
PR	05-MAR-1999;	PR	15-JUL-1999;	99US-0144005.
PR	09-MAR-1999;	PR	16-JUL-1999;	99US-0144085.
PR	23-MAR-1999;	PR	16-JUL-1999;	99US-0144086.
PR	25-MAR-1999;	PR	19-JUL-1999;	99US-0144325.
PR	29-MAR-1999;	PR	19-JUL-1999;	99US-0144331.
PR	01-APR-1999;	PR	19-JUL-1999;	99US-0144332.
PR	06-APR-1999;	PR	19-JUL-1999;	99US-0144333.
PR	08-APR-1999;	PR	19-JUL-1999;	99US-0144334.
PR	16-APR-1999;	PR	19-JUL-1999;	99US-0144335.
PR	19-APR-1999;	PR	20-JUL-1999;	99US-0144352.
PR	21-APR-1999;	PR	20-JUL-1999;	99US-0144632.
PR	23-APR-1999;	PR	20-JUL-1999;	99US-0144632.
PR	28-APR-1999;	PR	21-JUL-1999;	99US-0144884.
PR	30-APR-1999;	PR	21-JUL-1999;	99US-0144814.
PR	30-APR-1999;	PR	21-JUL-1999;	99US-0145086.
PR	04-MAY-1999;	PR	21-JUL-1999;	99US-0145088.
PR	04-MAY-1999;	PR	22-JUL-1999;	99US-0145085.
PR	05-MAY-1999;	PR	22-JUL-1999;	99US-0145087.
PR	06-MAY-1999;	PR	22-JUL-1999;	99US-0145089.
PR	06-MAY-1999;	PR	22-JUL-1999;	99US-0145192.
PR	07-MAY-1999;	PR	23-JUL-1999;	99US-0145145.
PR	11-MAY-1999;	PR	23-JUL-1999;	99US-0145218.
PR	14-MAY-1999;	PR	23-JUL-1999;	99US-0145224.
PR	14-MAY-1999;	PR	26-JUL-1999;	99US-0145276.
PR	14-MAY-1999;	PR	27-JUL-1999;	99US-0145913.
PR	14-MAY-1999;	PR	27-JUL-1999;	99US-0145918.
PR	18-MAY-1999;	PR	27-JUL-1999;	99US-0145919.
PR	19-MAY-1999;	PR	28-JUL-1999;	99US-0145951.
PR	20-MAY-1999;	PR	02-AUG-1999;	99US-0146386.
PR	21-MAY-1999;	PR	02-AUG-1999;	99US-0146388.
PR	24-MAY-1999;	PR	02-AUG-1999;	99US-0146389.
PR	25-MAY-1999;	PR	03-AUG-1999;	99US-0147038.
PR	27-MAY-1999;	PR	04-AUG-1999;	99US-0147204.
PR	28-MAY-1999;	PR	04-AUG-1999;	99US-0147302.
PR	01-JUN-1999;	PR	05-AUG-1999;	99US-0147192.
PR	03-JUN-1999;	PR	05-AUG-1999;	99US-0147260.
PR	04-JUN-1999;	PR	06-AUG-1999;	99US-0147303.
PR	07-JUN-1999;	PR	06-AUG-1999;	99US-0147416.
PR	08-JUN-1999;	PR	09-AUG-1999;	99US-0147493.
PR	10-JUN-1999;	PR	09-AUG-1999;	99US-0147935.
PR	10-JUN-1999;	PR	10-AUG-1999;	99US-0148171.
PR	14-JUN-1999;	PR	11-AUG-1999;	99US-0148319.
PR	16-JUN-1999;	PR	12-AUG-1999;	99US-0148341.
PR	16-JUN-1999;	PR	13-AUG-1999;	99US-0148565.
PR	17-JUN-1999;	PR	13-AUG-1999;	99US-0148684.
PR	18-JUN-1999;	PR	16-AUG-1999;	99US-0149368.
PR	18-JUN-1999;	PR	17-AUG-1999;	99US-0149175.
PR	18-JUN-1999;	PR	18-AUG-1999;	99US-0149426.
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149722.
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149929.
PR	18-JUN-1999;	PR	23-AUG-1999;	99US-0149902.
PR	18-JUN-1999;	PR	23-AUG-1999;	99US-0149930.
PR	18-JUN-1999;	PR	25-AUG-1999;	99US-0150566.
PR	18-JUN-1999;	PR	26-AUG-1999;	99US-0150884.
PR	18-JUN-1999;	PR	27-AUG-1999;	99US-0151065.
PR	21-JUN-1999;	PR	27-AUG-1999;	99US-0151066.
PR	22-JUN-1999;	PR	30-AUG-1999;	99US-0151303.
PR	23-JUN-1999;	PR	31-AUG-1999;	99US-0151438.
PR	23-JUN-1999;	PR	01-SEP-1999;	99US-0151930.
PR	24-JUN-1999;	PR	07-SEP-1999;	99US-0152363.
PR	28-JUN-1999;	PR	10-SEP-1999;	99US-0153070.
PR	29-JUN-1999;	PR	13-SEP-1999;	99US-0153758.
PR	30-JUN-1999;	PR	15-SEP-1999;	99US-0154018.
PR	01-JUL-1999;	PR	16-SEP-1999;	99US-0154039.
PR		PR	20-SEP-1999;	99US-0154779.

```
PR 22-SEP-1999; 99US-01551139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 3.72e-42
Score: 816.00
Percent Similarity: 77.87%
Best Local Similarity: 65.16%
Query Match: 44.69%
DB: 21
Length: 1068
Matches: 159
Conservative: 31
Mismatch: 44
Indels: 10
Gaps: 3
```

US-09-720-384A-4 (1-343) x AAC35266 (1-1068)

```
QY 100 GlyAlaArgThrHisCysHisArgGlyIleGlyArgTrpValArgArgArgArgArgAsn 119
DB 220 GATGCTCAAACTCTGATGTCATACCAAAATGGATCTATCTCGAGGTAAATCCATTAAAC 279
QY 120 GlyAlaAlaProGlyGluAlaProHisSerProValLysGluLysProValMetSerAsn 139
DB 280 GGT-----CACAGGCAAAAGCAGGACCT---TTGTCTACG 315
QY 140 IleGlyLysSerThrAsnIleLeuTrpHisAsnCysLeuIleGlyGlnSerAspArgGln 159
DB 316 CTCGGAAACTCGACAAATATAAGTGGCATGAATGTTCTGTTCGAGAAAGTTGATAGACAG 375
QY 160 LysLeuLeuGlyGlnLysGlyCysValValTrpIleThrGlyLeuSerClySerClyLys 179
DB 376 AGATTGCTTGTATCAGAAAGGATGCTGATTGGGTACCCGCTTAGTGGTTTCAGGGAG 435
QY 180 SerThrLeuAlaCysAlaLeuSerArgGlyLeuHisCysArgGlyHisLeuThrTyrVal 199
DB 436 ACTACTTGGCTGTGCTTGTGAATCAGATGTTGTATCAAAAGGGAAGCTTGTATAT 495
QY 200 LeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPhelysAlaGluAsp 219
DB 496 CTTCATGCTGATATGCTTAGGCGATGCTTAACCGCGATCTTAGCTTTAAAGCTGAGGAT 555
```

```
QY 220 ArgAlaGluAsnIleArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyVal 239
DB 556 CGTGCAGAGAATATTCGTAGAGTTGGAGAGTTGCTAAGCTTTTTCGGGATGCTGGAATA 615
QY 240 IleCysIleAlaSerLeuIleSerProTyrArgArgAspArgAspAlaCysArgAlaLeu 259
DB 616 ATTCGATTGCGAGTTTGTATCTCTTATAGAACAGATAGGACGCTTGTCTGAAGTTTG 675
QY 260 LeuProHisSerAspPheIleGluValPheIleAspLeuProLeuLysIleCysGluAla 279
DB 676 CTCGCCAGGAGATTTTGTAGGTGTTCATGATGTACCCCTTAGTGTTCGAGCGC 735
QY 280 ArgAspProLysGlyLeuTyrLysLeuAlaArgThrGlyLysIleLysGlyPheThrGly 299
DB 736 AGGGATCCAAAGGCTTTTACAAGCTTCTCGTCGAGAAAGATCAAAAGTTTATCCGGG 795
QY 300 IleAspAspProTyrGluProIleAsnGlyGluIleValIleLysMetLysAspGlu 319
DB 796 ATCGATGACCTTAGGACCCACCATTTGAATGCGAG-----ATTCTCTAGGACGTGAA 849
QY 320 GluCysProSerProLysAlaMetAlaLysGlnValLeuCysTyrLeuGluGluAsnGly 339
DB 850 GGAGGAACCTTCTCTATCGAAATGCGGAAAGGTCTCGGATAGTACTTAGATAACAAGGT 909
QY 340 TyrLeuGlnAla 343
DB 910 TATCTTCAAGCA 921
RESULT 8
AAC45659
ID AAC45659 standard; DNA; 1183 BP.
XX
AC AAC45659;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 47294.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126284.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
```

PR 11-MAY-1999;	99US-0134256.	PR 23-JUL-1999;	99US-0145218.
PR 14-MAY-1999;	99US-0134218.	PR 23-JUL-1999;	99US-0145224.
PR 14-MAY-1999;	99US-0134219.	PR 26-JUL-1999;	99US-0145276.
PR 14-MAY-1999;	99US-0134221.	PR 27-JUL-1999;	99US-0145913.
PR 18-MAY-1999;	99US-0134370.	PR 27-JUL-1999;	99US-0145918.
PR 19-MAY-1999;	99US-0134768.	PR 27-JUL-1999;	99US-0145919.
PR 20-MAY-1999;	99US-0134941.	PR 28-JUL-1999;	99US-0145951.
PR 21-MAY-1999;	99US-0135124.	PR 02-AUG-1999;	99US-0146386.
PR 21-MAY-1999;	99US-0135353.	PR 02-AUG-1999;	99US-0146388.
PR 24-MAY-1999;	99US-0135629.	PR 02-AUG-1999;	99US-0146389.
PR 25-MAY-1999;	99US-0136021.	PR 03-AUG-1999;	99US-0147038.
PR 27-MAY-1999;	99US-0136292.	PR 04-AUG-1999;	99US-0147204.
PR 28-MAY-1999;	99US-0136782.	PR 04-AUG-1999;	99US-0147302.
PR 01-JUN-1999;	99US-0137222.	PR 05-AUG-1999;	99US-0147192.
PR 03-JUN-1999;	99US-0137528.	PR 05-AUG-1999;	99US-0147260.
PR 04-JUN-1999;	99US-0137502.	PR 06-AUG-1999;	99US-0147303.
PR 07-JUN-1999;	99US-0137724.	PR 06-AUG-1999;	99US-0147416.
PR 08-JUN-1999;	99US-0138094.	PR 09-AUG-1999;	99US-0147493.
PR 10-JUN-1999;	99US-0138540.	PR 09-AUG-1999;	99US-0147935.
PR 10-JUN-1999;	99US-0138847.	PR 10-AUG-1999;	99US-0148171.
PR 14-JUN-1999;	99US-0139119.	PR 11-AUG-1999;	99US-0148319.
PR 16-JUN-1999;	99US-0139452.	PR 12-AUG-1999;	99US-0148341.
PR 16-JUN-1999;	99US-0139453.	PR 13-AUG-1999;	99US-0148565.
PR 17-JUN-1999;	99US-0139492.	PR 13-AUG-1999;	99US-0148684.
PR 18-JUN-1999;	99US-0139454.	PR 16-AUG-1999;	99US-0149368.
PR 18-JUN-1999;	99US-0139455.	PR 17-AUG-1999;	99US-0149175.
PR 18-JUN-1999;	99US-0139456.	PR 18-AUG-1999;	99US-0149426.
PR 18-JUN-1999;	99US-0139457.	PR 20-AUG-1999;	99US-0149722.
PR 18-JUN-1999;	99US-0139458.	PR 20-AUG-1999;	99US-0149723.
PR 18-JUN-1999;	99US-0139459.	PR 20-AUG-1999;	99US-0149929.
PR 18-JUN-1999;	99US-0139460.	PR 23-AUG-1999;	99US-0149902.
PR 18-JUN-1999;	99US-0139461.	PR 23-AUG-1999;	99US-0149930.
PR 18-JUN-1999;	99US-0139462.	PR 25-AUG-1999;	99US-0150566.
PR 18-JUN-1999;	99US-0139463.	PR 26-AUG-1999;	99US-0150884.
PR 18-JUN-1999;	99US-0139750.	PR 27-AUG-1999;	99US-0151065.
PR 18-JUN-1999;	99US-0139763.	PR 27-AUG-1999;	99US-0151066.
PR 21-JUN-1999;	99US-0139817.	PR 27-AUG-1999;	99US-0151080.
PR 22-JUN-1999;	99US-0139899.	PR 30-AUG-1999;	99US-0151303.
PR 23-JUN-1999;	99US-0140353.	PR 31-AUG-1999;	99US-0151438.
PR 23-JUN-1999;	99US-0140354.	PR 01-SEP-1999;	99US-0151930.
PR 24-JUN-1999;	99US-0140695.	PR 07-SEP-1999;	99US-0152363.
PR 28-JUN-1999;	99US-0140823.	PR 10-SEP-1999;	99US-0153070.
PR 29-JUN-1999;	99US-0140891.	PR 13-SEP-1999;	99US-0153758.
PR 30-JUN-1999;	99US-0141287.	PR 15-SEP-1999;	99US-0154018.
PR 01-JUL-1999;	99US-0141842.	PR 16-SEP-1999;	99US-0154039.
PR 01-JUL-1999;	99US-0142154.	PR 20-SEP-1999;	99US-0154779.
PR 02-JUL-1999;	99US-0142055.	PR 22-SEP-1999;	99US-0155139.
PR 06-JUL-1999;	99US-0142390.	PR 23-SEP-1999;	99US-0155486.
PR 08-JUL-1999;	99US-0142803.	PR 24-SEP-1999;	99US-0155659.
PR 09-JUL-1999;	99US-0142920.	PR 28-SEP-1999;	99US-0156458.
PR 12-JUL-1999;	99US-0142977.	PR 29-SEP-1999;	99US-0156596.
PR 13-JUL-1999;	99US-0143542.	PR 04-OCT-1999;	99US-0157117.
PR 14-JUL-1999;	99US-0143624.	PR 05-OCT-1999;	99US-0157753.
PR 15-JUL-1999;	99US-0144005.	PR 06-OCT-1999;	99US-0157865.
PR 16-JUL-1999;	99US-0144085.	PR 07-OCT-1999;	99US-0158029.
PR 16-JUL-1999;	99US-0144086.	PR 08-OCT-1999;	99US-0158232.
PR 19-JUL-1999;	99US-0144325.	PR 12-OCT-1999;	99US-0158369.
PR 19-JUL-1999;	99US-0144331.	PR 13-OCT-1999;	99US-0159293.
PR 19-JUL-1999;	99US-0144332.	PR 13-OCT-1999;	99US-0159294.
PR 19-JUL-1999;	99US-0144333.	PR 13-OCT-1999;	99US-0159295.
PR 19-JUL-1999;	99US-0144334.	PR 14-OCT-1999;	99US-0159329.
PR 19-JUL-1999;	99US-0144335.	PR 14-OCT-1999;	99US-0159330.
PR 20-JUL-1999;	99US-0144352.	PR 14-OCT-1999;	99US-0159331.
PR 20-JUL-1999;	99US-0144632.	PR 14-OCT-1999;	99US-0159637.
PR 21-JUL-1999;	99US-0144884.	PR 18-OCT-1999;	99US-0159638.
PR 21-JUL-1999;	99US-0144814.	PR 18-OCT-1999;	99US-0159584.
PR 21-JUL-1999;	99US-0145086.	PR 21-OCT-1999;	99US-0160741.
PR 21-JUL-1999;	99US-0145088.	PR 21-OCT-1999;	99US-0160767.
PR 22-JUL-1999;	99US-0145085.	PR 21-OCT-1999;	99US-0160768.
PR 22-JUL-1999;	99US-0145087.	PR 21-OCT-1999;	99US-0160770.
PR 22-JUL-1999;	99US-0145089.	PR 21-OCT-1999;	99US-0160814.
PR 22-JUL-1999;	99US-0145192.	PR 21-OCT-1999;	99US-0160815.
PR 23-JUL-1999;	99US-0145145.	PR 22-OCT-1999;	99US-0160980.

PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:

Pred. No.: 4,15e-42 Length: 1183
Score: 816.00 Matches: 159
Percent Similarity: 77.87% Conservative: 31
Best Local Similarity: 65.16% Mismatches: 44
Query Match: 44.69% Indels: 10
DB: 21 Gaps: 3

US-09-720-384A-4 (1-343) x AAC45659 (1-1183)

QY 100 GlyAlaArgThrHisCysHisArgGlyIleGlyArgTrpValArgArgArgArgGln 119
|||||:||||| ||| ||| :|||
Db 221 GGATCTCAAACTCGAGTCATACAAAATGATCTATTCCTGAGGTAAATCCATTAAAC 280
QY 120 GlyAlaAlaProGlyGluAlaProHisSerProValLysGluLysProValMetSerAsn 139
|||||:||||| ||| ||| :|||
Db 281 GGT-----CACACGGGACAAAAGGACCT--TTGCTACG 316
QY 140 IleGlyLysSerThrAsnIleLeuTrpHisAsnCysLeuIleGlyGlnSerAspArgGln 159
|||||:||||| ||| ||| ||| :|||
Db 317 GTCGAAACTCGCAAAATATAAAGTGGCATGAATGTTCTGTGAGAAAGTTGATAGACAG 376
QY 160 LysLeuLeuGlyGlnLysGlyCysValValPleThrGlyLeuSerGlySerGlyLys 179
|||||:||||| ||| ||| ||| ||| :|||
Db 377 AGATTGCTTGATCAGAAAGGATGTGTGATTTGGGTACCGGCTCTTAGTGGTTCAGGGAAG 436
QY 180 SerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTyrVal 199
|||||:||||| ||| ||| ||| ||| :|||
Db 437 AGTACTTTGGCTTGTGCTTTGAATCAGATGTTGTATCAAAAGGGAAGCTTTGTTATATT 496
QY 200 LeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLysAlaGluAsp 219
|||||:||||| ||| ||| ||| ||| ||| :|||
Db 497 CTTGATGTGTATATGTAGGCATGGCTTAAACCGTGATCTTAGCTTTAAAGCTGAGGAT 556
QY 220 ArgAlaGluAsnIleArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyVal 239
|||||:||||| ||| ||| ||| ||| ||| :|||
Db 557 CGTGACAGAAATATTCGTAGAGTTGAGAGGTTGCTAAGCTTTTTCGGGATGCTGGAATA 616
QY 240 IleCysIleAlaSerLeuIleSerProTyrArgArgAspArgAspAlaCysArgAlaLeu 259
|||||:||||| ||| ||| ||| ||| ||| ||| :|||
Db 617 ATCTGCATTCGCGATTTGTATATCTCCCTTATAGACAGATAGGACGCTTGTCCAAGTTTG 676
QY 260 LeuProHisSerAsnPheIleGluValPheIleAspLeuProLeuLysIleCysGluAla 279
|||||:||||| ||| ||| ||| ||| ||| ||| :|||
Db 677 CTCCTCCGAGGAGATTTTGTAGGTGTTCATGGATGTACCGCTTAGTCTTTCGAGGCG 736
QY 280 ArgAspProLysGlyLeuTrpLysLeuAlaArgThrGlyLysIleLysGlyPheThrGly 299
|||||:||||| ||| ||| ||| ||| ||| ||| ||| :|||
Db 737 AGGATCAAAAGGTCCTTACAGCTTCTGTCAGAGAAATCAAGGTTTATACCGGG 796
QY 300 IleAspAspProTyrGluProProIleAsnGlyGluIleValIleLysMetLysAspGlu 319
|||||:||||| ||| ||| ||| ||| ||| ||| ||| :|||
Db 797 ATCGATGACCTTACGAGCCACCATTAAGTCCGAG-----ATTCTCTAGGACGTGAA 850
QY 320 GluCysProSerProLysAlaMetAlaLysGlnValLeuCysTyrLeuGluAsnGly 339
|||||:||||| ||| ||| ||| ||| ||| ||| ||| :|||
Db 851 GGAGGAACCTTCTCTCCATCGAATGCGGAAAGGTCGCGATACCTAGTAAACAGGCT 910
QY 340 TyrLeuGlnAla 343
|||||:||||| ||| ||| ||| ||| ||| ||| ||| :|||

Db 911 TATCTTCAAGCA 922
RESULT 9
AAC51190
ID AAC51190 standard; DNA; 714 BP.
XX AAC51190; AC
XX AAC51190; AC
XX 18-OCT-2000 (first entry)
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 67599.
XX Arabidopsis thaliana.
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
OS Arabidopsis thaliana.
PN EP1033405-A2.
XX 06-SEP-2000.
PD 25-FEB-2000; 2000EP-0301439.
PF 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126284.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.


```
QY 243 aserLeulleSerProTyrhrghrghspargaspAlaCysArgAlaLeuLeuProHisSe 263
|||||
Db 601 AAGTTTGATTTCTCCGTACCGGAGACAGACGCGTCCGTCTTACCTGACGG 660
|||||
QY 263 rsnPheileGluValPheileAspLeuProLeuLysileCysGluAlaArgAspProLy 283
|||||
Db 661 CGATTTCGTGAGGTCTTCATGGAGCTTCCTTCATGTCGTGGAGTCGAGATCCAAA 720
|||||
QY 283 sglYLeuTyrLysLeuAlaArgThrGlyLysileLysGlyPheThrGlyLysAspPr 303
|||||
Db 721 GGGGTTGTACAAAGCTTCACGTCAGGCAAAATCAAAAGGTTCACTGGAATCGACGCC 780
|||||
QY 303 oTyrGluProProileAsnGlyIleValileLys---MetLysAspGluGluCysPr 322
|||||
Db 781 TTACGAGCCACCAGTGAATTCGAGGTAGTGTGAACACACAGGAGACGAGTCGTG 840
|||||
QY 322 oSerProLysAlaMetAlaLysGlnValLeuCysTyrLeuGluGluAsnGlyTyrLeuGl 342
|||||
Db 841 TTGGCCACGTGAGATGCTGAGAACATCATCTCTTACCTGCAAAACAAAGGTTATCTTGA 900
|||||
QY 342 n 342
;
Db 901 G 901

RESULT 11
AAC42294
ID AAC42294 standard; DNA; 948 BP.
XX AC AAC42294;
XX XX
XX DT 17-OCT-2000 (first entry)
XX DE
XX XX
XX KW Arabidopsis thaliana DNA fragment SEQ ID NO: 35008.
XX KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; terminator; termination sequence; ss.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD
XX PF 06-SEP-2000.
XX XX
XX XX 25-FEB-2000; 2000EP-0301439.
XX XX
XX XX 25-FEB-1999; 99US-0121825.
XX XX 05-MAR-1999; 99US-0123180.
XX XX 09-MAR-1999; 99US-0123548.
XX XX 23-MAR-1999; 99US-0125788.
XX XX 23-MAR-1999; 99US-0126264.
XX XX 29-MAR-1999; 99US-0126785.
XX XX 01-APR-1999; 99US-0127462.
XX XX 06-APR-1999; 99US-0128234.
XX XX 08-APR-1999; 99US-0128714.
XX XX 16-APR-1999; 99US-0129845.
XX XX 19-APR-1999; 99US-0130077.
XX XX 21-APR-1999; 99US-0130449.
XX XX 23-APR-1999; 99US-0130510.
XX XX 23-APR-1999; 99US-0130891.
XX XX 28-APR-1999; 99US-0131449.
XX XX 30-APR-1999; 99US-0132048.
XX XX 30-APR-1999; 99US-0132407.
XX XX 04-MAY-1999; 99US-0132484.
XX XX 05-MAY-1999; 99US-0132485.
XX XX 06-MAY-1999; 99US-0132486.
XX XX 06-MAY-1999; 99US-0132487.
XX XX 07-MAY-1999; 99US-0132863.
XX XX 11-MAY-1999; 99US-0134256.
XX XX 14-MAY-1999; 99US-0134218.
XX XX 14-MAY-1999; 99US-0134219.
XX XX 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
```

```
PR 27-JUL-1999; 99US-0145318.
PR 27-JUL-1999; 99US-0145319.
PR 28-JUL-1999; 99US-0145351.
PR 28-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 12-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
```

```
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 28-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 5.84e-40 Length: 948
Score: 778.50 Matches: 175
Percent Similarity: 66.04% Conservative: 37
Best Local Similarity: 54.52% Mismatches: 67
Query Match: 42.63% Indels: 43
DB: 21 Gaps: 9

US-09-720-384A-4 (1-343) x AAC42294 (1-948)

Qy 40 ProThrProThrLeuAlaValIleLeuValAsnProGlnArgAlaProProVal----- 57
Db 55 CCCTCGGATTCAATGGAAGGATTAGTATTCAGAGCATCTCGA---CCGTCCGTTTCTGT 111
Qy 58 ---LeuProGlyLeu-----ThrProSerAspAla-----Pro 67
Db 112 TCTATTCCAGGTCTCGCGCGGATTCCACCGAAACCTCCAAGTGACGGTTTCTCAAG 171
Qy 68 LeuProAlaLeuValIleHisGlyLeuThrProArgSerSerHisSerAlaGlyLeu 87
Db 172 CTGCTCGGCTC-GTCTATTCGCGGACACCGGAAATAGTAGGAAATCTACTTCCCTT 230
Qy 88 AlaSerAspSerGlyArgArgGluGlyArgGlyAlaArgThrHisCysHisArg 107
Db 231 TCATCCAATCTCAGCCGT-----TAACGTCTCTGCTCAAGTCTCCCTCACC 278
Qy 108 GlyLeGlyArg-----TrpValArgArg-ArgArgArgAsnGlyAlaAlaPr 123
Db 279 TGATTTTCCGCCCTTCAGAACTATATCTGAAGAGGAGAAATACCGA----- 330
Qy 123 oGlyGluAlaProHisSerProValLysGluLysProValMetSerAsnIleGlyLysSe 143
Db 331 -----AAAGAG-----AAAGC 341
Qy 143 rThrAsnIleLeuTrpHisAsnCysLeuIleGlyGlnSerAspArgGlnLysLeuLeuG 163
Db 342 AGAGAACATCGTGTGGCAGAGATTCGATATCCAGATCCGACACACAACTTCTTCA 401
Qy 163 yGlnLysGlyCysValValTrpIleThrGlyLeuSerGlySerGlyLysSerThrLeuAl 183
Db 402 ACAAAGGGTGTGTCTGTTGGATCTACGTGCTCAGTGTTCAGGGAAGACACTGTTCG 461
Qy 183 aCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTyrValLeuAspGlyAs 203
Db 462 TGTGCACTAAGTAAGCACTTGTGAAAGAGGCAAACTTACTTACACACTCGACGCCA 521
Qy 203 pAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLysAlaGluAspArgAlaGluAs 223
Db 522 CAATGTCCGTACCGGCTTAACCGTGACCTCTTCAAGCAGAGCACCGCACCGAA 581
Qy 223 nIleArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyValIleCysIleAl 243
Db 582 CATTAGAAGATTCGTGAGGTGGCTAAGTTGTTCTGACGTCGGAGTCATTGTATACG 641
Qy 243 aSerLeuIleSerProTyrArgArgAspArgAspAlaCysArgAlaLeuProHisSe 263
Db 642 AAGTTTGATTTCTCCGTACCGGAGACAGACGCGTCCGCTCTGTGTACCTGACCG 701
Qy 263 rAsnPheIleGluValPheIleAspLeuProLeuLysIleCysGluAlaArgAspProLy 283
Db 702 CGATTTCTCGAGGTCTTCATGGACGTTCTCTTCATGTGTGCGAGTCGAGATCCAA 761
Qy 283 sGlyLeuTyrLysLeuAlaArgThrGlyLysIleLysGlyPheThrGlyIleAspAspPr 303
```

```
Db 762 GGGGTTGTACAAGCTTCAGCGTCAGGCAAAATCAAAAGGCTTCACCTCGAATCGACGACC 821
QY 303 OlyrGluProProlleasnGlyGluIleValIleLys---MetIlyAspGluGluCysPr 322
      ||||| |||:|||| |||:||||:|||| |||:||||
Db 822 TTACGAGCCACCACTGAATTGCCAGGTAGTGTGAACACACAGAGACGACGACGTCTGTG 881
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 322 oSerProLysAlaMetAlaLysGlnValIleuCystyrIeuGluGluAsnGlyTyrLeuG1 342
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 882 TTGCGCCAGTCAGATGCTGAGACATCATCTCTACCTGCAAAACAAGGTTATCTTGA 941
QY 342 n 342
Db 942 G 942

RESULT 12
AAC49070
ID AAC49070 standard; DNA; 1130 BP.
AC AAC49070;
XX
XX 18-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 59815.
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
XX Arabidopsis thaliana.
EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 29-APR-1999; 99US-0130891.
XX 30-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 30-APR-1999; 99US-0132407.
XX 04-MAY-1999; 99US-0132484.
XX 05-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 06-MAY-1999; 99US-0132487.
XX 07-MAY-1999; 99US-0132487.
XX 07-MAY-1999; 99US-0132863.
XX 11-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
XX 14-MAY-1999; 99US-0134370.
XX 18-MAY-1999; 99US-0134768.
XX 19-MAY-1999; 99US-0134941.
XX 20-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135353.
XX 24-MAY-1999; 99US-0135629.
XX 25-MAY-1999; 99US-0136021.
XX 27-MAY-1999; 99US-0136392.
XX 28-MAY-1999; 99US-0136782.
XX 01-JUN-1999; 99US-0137222.
XX 03-JUN-1999; 99US-0137528.
XX
XX 04-JUN-1999; 99US-0137502.
XX 07-JUN-1999; 99US-0137724.
XX 08-JUN-1999; 99US-0138094.
XX 10-JUN-1999; 99US-0138540.
XX 10-JUN-1999; 99US-0138847.
XX 14-JUN-1999; 99US-0139119.
XX 16-JUN-1999; 99US-0139452.
XX 16-JUN-1999; 99US-0139453.
XX 17-JUN-1999; 99US-0139492.
XX 18-JUN-1999; 99US-0139454.
XX 18-JUN-1999; 99US-0139455.
XX 18-JUN-1999; 99US-0139456.
XX 18-JUN-1999; 99US-0139457.
XX 18-JUN-1999; 99US-0139458.
XX 18-JUN-1999; 99US-0139459.
XX 18-JUN-1999; 99US-0139460.
XX 18-JUN-1999; 99US-0139461.
XX 18-JUN-1999; 99US-0139462.
XX 18-JUN-1999; 99US-0139463.
XX 18-JUN-1999; 99US-0139750.
XX 18-JUN-1999; 99US-0139763.
XX 21-JUN-1999; 99US-0139817.
XX 22-JUN-1999; 99US-0139899.
XX 23-JUN-1999; 99US-0140353.
XX 23-JUN-1999; 99US-0140354.
XX 24-JUN-1999; 99US-0140695.
XX 28-JUN-1999; 99US-0140823.
XX 29-JUN-1999; 99US-0140991.
XX 30-JUN-1999; 99US-0141287.
XX 01-JUL-1999; 99US-0141842.
XX 01-JUL-1999; 99US-0142154.
XX 02-JUL-1999; 99US-0142055.
XX 06-JUL-1999; 99US-0142390.
XX 08-JUL-1999; 99US-0142803.
XX 09-JUL-1999; 99US-0142920.
XX 12-JUL-1999; 99US-0142977.
XX 13-JUL-1999; 99US-0143542.
XX 14-JUL-1999; 99US-0143624.
XX 15-JUL-1999; 99US-0144005.
XX 16-JUL-1999; 99US-0144085.
XX 16-JUL-1999; 99US-0144086.
XX 19-JUL-1999; 99US-0144325.
XX 19-JUL-1999; 99US-0144331.
XX 19-JUL-1999; 99US-0144332.
XX 19-JUL-1999; 99US-0144333.
XX 19-JUL-1999; 99US-0144334.
XX 19-JUL-1999; 99US-0144335.
XX 20-JUL-1999; 99US-0144352.
XX 20-JUL-1999; 99US-0144632.
XX 20-JUL-1999; 99US-0144684.
XX 21-JUL-1999; 99US-0144814.
XX 21-JUL-1999; 99US-0145086.
XX 21-JUL-1999; 99US-0145088.
XX 22-JUL-1999; 99US-0145085.
XX 22-JUL-1999; 99US-0145087.
XX 22-JUL-1999; 99US-0145089.
XX 22-JUL-1999; 99US-0145192.
XX 23-JUL-1999; 99US-0145145.
XX 23-JUL-1999; 99US-0145218.
XX 23-JUL-1999; 99US-0145224.
XX 26-JUL-1999; 99US-0145276.
XX 27-JUL-1999; 99US-0145913.
XX 27-JUL-1999; 99US-0145918.
XX 27-JUL-1999; 99US-0145919.
XX 28-JUL-1999; 99US-0145951.
XX 02-AUG-1999; 99US-0146386.
XX 02-AUG-1999; 99US-0146388.
XX 02-AUG-1999; 99US-0146389.
XX 03-AUG-1999; 99US-0147038.
XX 04-AUG-1999; 99US-0147204.
XX 04-AUG-1999; 99US-0147302.
XX 05-AUG-1999; 99US-0147192.
XX 05-AUG-1999; 99US-0147260.
```


XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125798.
PR 25-MAR-1999; 99US-0126284.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128224.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145215.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.

PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159337.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:

Pred. No.: 3, 37e-38 Length: 1133
Score: 752.50 Matches: 143
Percent Similarity: 79.26% Conservative: 29
Best Local Similarity: 65.90% Mismatches: 30
Query Match: 41.21% Indels: 15
DB: 21 Gaps: 2

US-09-720-384A-4 (1-343) x AAC36627 (1-1133)

QY 141 GlyLysSerThrAsnIleLeuTrpHisAsnCysLeuIleGlyGlnSerAspArgGlnLys 160
Db 345 GGAAAGCAAAAGACATGTCTGGCATGATTGTCOCGTACTAAATCCGACAGGCAAGAA 404
QY 161 LeuLeuGlyGlnLysGlyCysValValTrpIleThrGlyLeuSerGlySerGlyLysSer 180
Db 405 TTAATTAAGCATAGGAGATGTGTGATTGGATTACTGGCTTAAGTGGTTACAGTAAAGT 464
QY 181 ThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTyrValLeu 200
Db 465 AGTCGGCATGTGCTTACTGAGCTTTGCACAATCGTGAAGCTTTTCGTATATACTT 524
QY 201 AspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLysAlaGluAspArg 220
Db 525 GATGGTGACATGTTCCACATGGTTTAACACCGGATCTAGTTTCGACGACATGATCGA 584
QY 221 AlaGluAsnIleArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyValIle 240
Db 585 GGTGAACATTCCTGCAAGAGTTGGTGAAGTGGCTAACTGTTTCGACATTTCTGGTATTATC 644
QY 241 CysIleAlaSerLeuIleSerProTyrArgArgAspAlaCysArgAlaLeuLeu 260
Db 645 TGTATTGAAGTTAATATCTCTACCGGATAGACGAGCTGCCTGCCGTGCATTATTA 704
QY 261 ProHisSerAsnPheIleGluValPheIleAspLeuProLeuLysIleCysGluAlaArg 280
Db 705 CCACAAGGAGATTTCAATGAGGTATTTATGGATGTGCCACTCCATGTTTGTGAAGCTAGA 764

QY 281 AspProLysGlyLeuTyrLysLeuAlaArgThrGlyLysIleLysGlyPheThrGlyIle 300
Db 765 GATCCAAAGGGCTTATACAAACGTCACGCGCTGGTAAATCAAAATCAAAAGGTTTTACAGGAGTA 824
QY 301 AspAspProTyrGluProProIleAsnGlyGluIleValIleLys----- 315
Db 825 GATGATCCATATGAAGCGCCTTTGGATTGGGAGATTGTAATACAAACAGTAGACAAG 884
QY 316 -----MetLysAspGluGluCysProSerProLysAla----- 326
Db 885 GGGCTTCTTCATCATCTTCATCTTCACCTTCATCTTCGCTCTTCTCTCTGTGT 944
QY 327 ---MetAlaLysGlnValLeuCystyrLeuGluGluAsnGlyTyrLeuGln 342
Db 945 GAAATGCGACAGATATTGTTGTGCTGCTACTTGGACCAAAATGGATACCTGAAG 995
RESULT 14
AAC51315
ID AAC51315 standard; DNA; 483 BP.
XX
AC AAC51315;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 68060.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.

PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139463.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 30-JUN-1999; 99US-0140991.
PR 01-JUL-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142354.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149436.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```
Alignment Scores:
Pred. No.: 3,04e-30 Length: 483
Score: 617.50 Matches: 117
Percent Similarity: 85.09% Conservative: 20
Best Local Similarity: 72.67% Mismatches: 15
Query Match: 33.82% Indels: 9
DB: 21 Gaps: 1

US-09-720-384A-4 (1-343) x AAS94855 (1-483)
QY 137 MetSerAsnIleGlyLysSerThrAsnIleLeuTrpHisAsnCysLeuIleGlyGlnSer 156
DB 1 ATGTGACAGTGGGAATTCACGACACATATTTGGCAAGAAATCCCCATTTGGGAACACT 60
QY 157 AspArgGlnLysLeuLeuGlyGlnLysGlyCysValValTrpIleThrGlyLeuSerGly 176
DB 61 GAAAGGCAGAAAGTTGTTAAATCAGAAAGGTTGTGTGTGGATTCACAGGGCTTCAGTGCC 120
QY 177 SerGlyLysSerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeu 196
DB 121 TCAGGGAACACGCTTAGCTTCGCTCTAGTAGAGAGCTGACACACCGGGGAAGCTA 180
QY 197 ThrTyrValLeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLys 216
DB 181 TCATATATTCTTGATGTTGATCAATCTTCGCTCATGTTGTTGAACAAAGATCTTGGTTTCAAG 240
QY 217 AlaGluAspArgAlaGluAsnIleArgValGlyGluValAlaLysLeuPheAlaAsp 236
DB 241 GCAGAGGATACAGTGGAAATATACACGAGGCTCGGAGAGTAGCAAACTCTTTGGCGAT 300
QY 237 AlaGlyValIleCysIleAlaSerLeuIleSerProTyrArgArgAspAlaCys 256
DB 301 GCTGGTTTAATCTGTATTGTCACGCTCATATCCCGTATAGAAAGACCGTGACGCTGC 360
QY 257 ArgAlaLeuLeuProHisSerAsnPheIleGlu----- 267
DB 361 AGGGAATATACAGAAATTCATCTTTATTAGGCAAGCTATTCAAAACCGCTTTACA 420
QY 268 ValPheIleAspLeuProLysIleCysGluAlaArgaspProLysGlyLeuTyrLys 287
DB 421 GTTTCATGAACATGCTCTGCAATTTGTGAAGCAAGGGACCCCTAAAGGCCTATACAAG 480
QY 288 Leu 288
DB 481 CTT 483

RESULT 15
AAS94855
ID AAS94855 standard; DNA; 2646 BP.
AC AAS94855;
XX
XX 14-FEB-2002 (first entry)
XX
XX Human DNA sequence #110 expressed during foam cell differentiation.
XX
XX Human; foam cell differentiation; atherosclerosis; cerebral stroke;
XX
XX cardiovascular disorder; coronary artery disease; gene therapy; ds.
XX
XX Homo sapiens.
XX
XX WO200177389-A2.
XX
XX 18-OCT-2001.
XX
XX 04-APR-2001; 2001WO-US11128.
XX
XX 05-APR-2000; 2000US-195106P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Shiffman D, Somogyi R, Lawn R, Seilhamer JJ, Porter GJ, Mikita T;
```

PI Tai J;

```
XX WPI; 2002-010925/01.
XX
XX Composition useful for diagnosis of conditions, disorders or diseases
XX associated with atherosclerosis, comprises several polynucleotides that
XX are differentially expressed in foam cell development -
XX
XX Claim 1; Page 162-163; 315pp; English.
XX
XX The present invention relates to the isolation of human polynucleotide
XX sequences that are differentially expressed during foam cell
XX differentiation. The polynucleotide sequences of the invention or a
XX composition comprising these polynucleotides are useful as a high
XX throughput method for detecting altered expression of one or more
XX polynucleotides in a sample. The polynucleotides can be used in the
XX diagnosis of disorders associated with foam cell development such as
XX atherosclerosis, cerebral stroke, and cardiovascular disorders such as
XX coronary artery disease. The polynucleotide sequences can also be used
XX as PCR primers and probes. The polynucleotides of the invention are also
XX useful in gene therapy. AAS94746-AAS95021 represent the human
XX polynucleotide sequences of the invention which are differentially
XX expressed during foam cell differentiation.
XX
XX Sequence 2646 BP; 715 A; 566 C; 634 G; 727 T; 4 other:
XX
XX Alignment Scores:
XX Pred. No.: 2,51e-22 Length: 2646
XX Score: 502.50 Matches: 123
XX Percent Similarity: 56.01% Conservative: 40
XX Best Local Similarity: 42.27% Mismatches: 84
XX Query Match: 27.52% Indels: 45
XX DB: 24 Gaps: 7
XX
XX US-09-720-384A-4 (1-343) x AAS94855 (1-2646)
XX
XX QY 64 SerAspAlaProLeuProAlaLeuValIleHisGlyLeuThrProArgSerSerHisSer 83
XX DB 58 AGCGCGCGCGCTACCCGACGCGAGGAGCGCCGCTCCAGCCCGCCGCTCGCTGG 117
XX
XX QY 84 SerAlaGlyLeuAlaSerAspSerGlyArgGlyGlyGlyGlyGlyGlyGlyGlyGly 103
XX DB 118 CCGTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 172
XX
XX QY 104 HisCysHisArgGlyIleGlyArgTrpValArgArgArgArgArgArgArgArgArg 123
XX DB 173 -----TCCGCGGTC 181
XX
XX QY 124 GlyGluAlaProHisSerProValLysGlyLysProValMetSerAsn----- 139
XX DB 182 ATGGAGATCCCGGGAGCTTGTGCAAGAAAGTCAAGTCAGGCAATACGCGCAGAACTGG 241
XX
XX QY 140 ---IleGlyLysSerThrAsnIleLeuTrpHisAsnCysLeuIleGlyGlnSerAspArg 158
XX DB 242 GGAATGCAGAGCAACCAATGTCACTACCAAGCCCATCATGTCCAGCAGCAACAAGAGA 301
XX
XX QY 159 GlnLysLeuLeuGlyGln-----LysGlyCysValValTrpIleThrGlyLeu 174
XX DB 302 GGTGAGTGGTGGGACCAAGAGGCGCTTCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGG 361
XX
XX QY 175 SerGlySerGlyLysSerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGly 194
XX DB 362 TCTGGAGCGGGAAGACACTACTGTGAGCATGGCCTTGGAGGAGTACCTGTTGTCTGGT 421
XX
XX QY 195 HisLeuThrTyrValLeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSer 214
XX DB 422 ATTCATGCTACACTCTGGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 481
XX
XX QY 215 PheLysAlaGluAspArgAlaGluAsnIleArgArgValGlyGlyValAlaLysLeuPhe 234
XX DB 482 TTTAGTCCCTGAACAGACAGAGAGAATGTTCCGACGATCCGAGAAAGTTGCTAAACTGTT 541
XX
XX QY 235 AlaAspAlaGlyValIleCysIleAlaSerLeuIleSerProTyrArgArgAspArgAsp 254
```

```

|||||
Db 542 GCAGATGCTGGCTTAGTGTGCATCACAAAGTTTCATATCACCTTACACTCAGGATCGCAAC 601
QY 255 AlaCysArg-----AlaLeuLeuProHisSerAsnPheIleGluValPhe 269
Db 602 AATGCAAGCGCAAAATTCATGAAGGTGCAAGTTTACCG-----TTTTTTGAAGTATTT 652
QY 270 IleAspLeuProLeuLysIleCysGluAlaArgAspProLysGlyLeuTyrLysLeuAla 289
Db 653 GTTGATGCTCCTCTCATGTGTTGTGAACACAGAGGGATGTCAAAGGACTCTACAAAAAGCC 712
QY 290 ArgThrGlyLysIleLysGlyPheThrGlyIleAspAspProTyrGluProProIleAsn 309
Db 713 CGGGCAGGAGAAATTAAGGTTTCACCTGGGATCGATTCTGAATATGAAAAAGCCAGAGGCC 772
QY 310 GlyGluIleValIleLysMetLysAsp-----GluGluCysProSerProLysAla 326
Db 773 CCTGAGTTGGTCTGAAAACACAGACTCTGTGATGTAAATGACTGT----- 817
QY 327 MetAlaLysGlnValLeuCysTyrLeuGluGlu 337
Db 818 ---GTCACAGCAAGTTGTGGAACCTTCTACAGGAA 847
```

Search completed: November 2, 2002, 03:08:48
Job time : 201 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 2, 2002, 01:57:13 : Search time 45 Seconds
(without alignments)
1831.572 Million cell updates/sec

Title: US-09-720-384A-4

Perfect score: 1826

Sequence: 1 RPFHINQTEPLVTHQOPP.....PKAMQVLCYLENGYLQA 343

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus_p2n_model -DEV=xlh
-O=/cgn2_1/USPTO_spool/US09720384/runat_30102002_090527_19203/app_query.fasta_1.519
-DB-Issued_Patents_NA -QMT=fastap -SUFFIX=rnl -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -ONITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09720384 -CGN_1_1_20_runat_30102002_090527_19203 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued_Patents_NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	496	27.2	2506	1 US-08-879-561-4	Sequence 4, Appli
2	475	26.0	1160	4 US-09-153-310-41	Sequence 41, Appl
3	155	8.5	6803	3 US-08-665-259-19	Sequence 19, Appl
4	155	8.5	6803	3 US-08-762-500-19	Sequence 19, Appl
C 5	123.5	6.8	33529	4 US-09-144-085-3	Sequence 3, Appli
6	123	6.7	36741	4 US-09-301-665-3	Sequence 3, Appli
7	120.5	6.6	1377	4 US-09-461-474-5	Sequence 5, Appli
C 8	120.5	6.6	2626	1 US-08-156-020-3	Sequence 3, Appli
C 9	120.5	6.6	2626	1 US-08-156-020-5	Sequence 5, Appli
C 10	120.5	6.6	2626	1 US-08-156-020-7	Sequence 7, Appli
C 11	120.5	6.6	2626	1 US-08-156-020-9	Sequence 9, Appli
C 12	119.5	6.5	36519	3 US-08-923-137-2	Sequence 2, Appli

C 13	119	6.5	2499	3 US-07-602-848E-1	Sequence 1, Appli
C 14	117	6.4	1647	1 US-08-073-384C-9	Sequence 9, Appli
C 15	117	6.4	1647	1 US-08-254-359A-9	Sequence 9, Appli
C 16	117	6.4	1647	1 US-08-483-043-9	Sequence 9, Appli
C 17	117	6.4	1647	1 US-08-481-238-9	Sequence 9, Appli
C 18	117	6.4	1647	2 US-08-471-066B-9	Sequence 9, Appli
C 19	117	6.4	1647	2 US-08-484-356-9	Sequence 9, Appli
C 20	117	6.4	1647	2 US-08-757-653-9	Sequence 9, Appli
C 21	117	6.4	1647	2 US-08-599-491-9	Sequence 9, Appli
C 22	117	6.4	1647	2 US-08-756-386-9	Sequence 9, Appli
C 23	117	6.4	1647	3 US-08-682-853A-9	Sequence 9, Appli
C 24	117	6.4	1647	3 US-08-823-516-9	Sequence 9, Appli
C 25	117	6.4	1647	3 US-08-759-038-9	Sequence 9, Appli
C 26	117	6.4	1647	3 US-08-758-314-9	Sequence 9, Appli
C 27	117	6.4	1647	4 US-09-350-309-9	Sequence 9, Appli
C 28	117	6.4	2088	1 US-08-073-384C-10	Sequence 10, Appl
C 29	117	6.4	2088	1 US-08-254-359A-10	Sequence 10, Appl
C 30	117	6.4	2088	1 US-08-483-043-10	Sequence 10, Appl
C 31	117	6.4	2088	1 US-08-481-238-10	Sequence 10, Appl
C 32	117	6.4	2088	2 US-08-471-066B-10	Sequence 10, Appl
C 33	117	6.4	2088	2 US-08-484-356-10	Sequence 10, Appl
C 34	117	6.4	2088	2 US-08-757-653-10	Sequence 10, Appl
C 35	117	6.4	2088	2 US-08-599-491-10	Sequence 10, Appl
C 36	117	6.4	2088	2 US-08-756-386-10	Sequence 10, Appl
C 37	117	6.4	2088	2 US-08-516-10	Sequence 10, Appl
C 38	117	6.4	2088	3 US-08-682-853A-10	Sequence 10, Appl
C 39	117	6.4	2088	3 US-08-759-038-10	Sequence 10, Appl
C 40	117	6.4	2088	3 US-08-758-314-10	Sequence 10, Appl
C 41	117	6.4	2088	4 US-09-350-309-10	Sequence 10, Appl
C 42	117	6.4	2499	1 US-07-977-434-1	Sequence 1, Appli
C 43	117	6.4	2499	1 US-08-458-819-1	Sequence 1, Appli
C 44	117	6.4	2499	4 US-09-587-856-1	Sequence 1, Appli
C 45	117	6.4	2499	4 US-09-777-537-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-08-879-561-4
; Sequence 4, Application US/08879561
; Patent No. 5817482

GENERAL INFORMATION:

APPLICANT: Bandman, Olga
APPLICANT: Hillman, Jennifer L.
APPLICANT: Hawkins, Phillip R.
APPLICANT: Guegler, Karl J.
APPLICANT: Corley, Neil C.

TITLE OF INVENTION: DISEASE RELATED NUCLEOTIDE KINASES

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/879,561

FILING DATE: Herewith

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0325 US

TELECOMMUNICATION INFORMATION:

APPLICANT: Connors, Timothy D.
APPLICANT: Dackowski, William R.
APPLICANT: Van Raay, Terence J.
*APPLICANT: Klinger, Katherine W.
TITLE OF INVENTION: NOVEL HUMAN CHROMOSOME 16 GENES,
TITLE OF INVENTION: COMPOSITIONS, METHODS OF MAKING AND USING SAME
NUMBER OF SEQUENCES: 83
CORRESPONDENCE ADDRESS:
ADDRESSER: GENZYME CORPORATION
STREET: One Mountain Road
CITY: Framingham
STATE: Massachusetts
COUNTRY: United States of America
ZIP: 01701
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/762,500
FILING DATE: 09-DEC-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/665,259
FILING DATE: 17-JUN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/10469
FILING DATE: 17-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Dugan, Deborah A.
REGISTRATION NUMBER: 37,315
REFERENCE/DOCKET NUMBER: IGS-9.3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (508) 872-8400
TELEFAX: (508) 872-5415
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 6803 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-762-500-19
Alignment Scores:
Pred. No.: 0.00236 Length: 6803
Score: 155.00 Matches: 75
Percent Similarity: 34.64% Conservative: 22
Best Local Similarity: 26.79% Mismatches: 86
Query Match: 8.49% Indels: 97
DB: 3 Gaps: 13
US-09-720-384A-4 (1-343) x US-08-762-500-19 (1-6803)
QY 14 ThrHisThrGlnGlnProProSerProAlaProGlyProAlaSerGlnGlyGlnArgGln 33
Db 2555 TCCCGGGCTCCAGAGCCCTCCCGCTGGCCAGCCCGGCTGTAATCTGCTTCTGAT 2614
QY 34 GlyAsnThrLeuLeuSerProThrProThrLeuAlaValIleLeuValAsnProGln-Ar 53
Db 2615 TCCAGCTCTGGATGAGGCC---CCTCCCTCCCTCCCTCCCTCCCGACCGAGCAGC 2671
QY 53 gAlaProValLeuProGlyLeuThrProSerAspAlaProLeuProAlaLeuValI 73
Db 2672 CCGCCCGGCTGGCGGGCTGGCTGCTGCGCCCGCCACCC--- 2720
QY 73 eHisGlyLeuThrProArgSerSerHisSerAlaGlyLeuAlaSerAspSerGlyAr 93
Db 2721 -----TCCGTGGACAGCTCGTCCGCCCTCGCTGCACCGCGGAGGAG 2761
QY 93 gArgGlu----- 95

Db 2762 CGCGCGCGCTGTCACCGAGCGCCCGCCGCCACCGCCCTTCCTCCGGAGCGCGGAGA 2821
QY 95 ----- 95
Db 2822 CTGCTCCGCGCGCCCTCGGTGGTGGAGCGCGGGTGGGGCTTCGCGGGCGG 2881
QY 96 -GlyGluGlyArgGlyAlaArgThrHisCysHisArgGly-----IleGlyArgTrpVa 113
Db 2882 AGGCACCGGAGCGGGCGGCGAGCC---TGTTCATCGCTCTAGGCCCGAGCGAGG---AC 2935
QY 113 LarArgArgArgArgAsnGlyAlaAlaProGlyGluAlaProHisSerProValLysG 133
Db 2936 CGGCCAACATCCCGCTGCTGTGGCGCGGGCGGTGCGG----- 2978
QY 133 uLysProValMetSerAsnIleGlyLys-----Se 143
Db 2979 ---CCGCTGCTCCACCTCTGGCGGGCTGGGGCGCGCGGGGCCCTGTTCCTCGGC 3034
QY 143 rThrAsnIleLeuTrpHisAsnCysLeuIleGlyGlnSerAspArgGlnLysLeuG 163
Db 3035 ATTGCGGGCTGTGG-----GCAGAGCGCGGAGAGGCTTCTTTT 3076
QY 163 yGlnLysGlyCysValVal-----TrpIleThrGlyLeuSerGlySerGlyLy 179
Db 3077 CCCAAGGCGAGCGTCTTGGGGCGCGGCGCACTGGCTGACC---CGCAGCGGCTCGGCCA 3133
QY 179 sSerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTy 199
Db 3134 TGCTGTGGTGGCC-----CTGGGGGCTGC 3157
QY 199 lLeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLysAlaGluAs 219
Db 3158 TGCTGAGCGG-----AGGCACGCTCTTTCGCGCGCTGAGTCTGGCGCGCGCGGC 3208
QY 219 pArgAlaGluAsnIleArgValGlyGluValAlaLysLeuPheAlaAspAlaGly 238
Db 3209 CCGCGGACCCCTGCCACGATGAGGGGTGCGCCCGCGGCTGCTGCCAGGACTGGT 3266
RESULT 5
US-09-144-085-3/c
; Sequence 3, Application US/09144085
; Patent No. 6280999
; GENERAL INFORMATION:
; APPLICANT: Gustafsson, Claes
; APPLICANT: Betlach, Mary C.
; APPLICANT: Ashley, Gary
; APPLICANT: Julien, Bryan
; APPLICANT: Ziermann, Rainer
; TITLE OF INVENTION: SORANGIUM POLYKETIDE SYNTHASES AND ENCODING DNA
; FILE REFERENCE: 30062-20020.20
; CURRENT APPLICATION NUMBER: US/09/144,085
; EARLIER FILING DATE: 1998-08-31
; EARLIER APPLICATION NUMBER: 09/010,809
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 33529
; TYPE: DNA
; ORGANISM: Sorangium cellulosum
US-09-144-085-3
Alignment Scores:
Pred. No.: 13.1 Length: 33529
Score: 123.50 Matches: 86
Percent Similarity: 37.42% Conservative: 33
Best Local Similarity: 27.04% Mismatches: 126
Query Match: 6.76% Indels: 73
DB: 4 Gaps: 16
US-09-720-384A-4 (1-343) x US-09-144-085-3 (1-33529)


```
QY 18 GlnProSerProAla-----ProGlyProAlaSerGlnGlyGln 31
Db 13969 CGACCCCTGACCCAGCTCCCTGTCGCCGCGGTACACAGACGTCGCCGCGAGGTCCA 13910
QY 32 ArgGlnGlyAsnThrLeuLeuSerProThrProThrLeuAlaValIleLeuValAsnPro 51
Db 13909 GCTCGGGGCTCGGTGAGCTCTTCGCTACCGCTGTGCGCGCACACGCGCGCGGCCA 13850
QY 52 GlnArgAlaProVal-----LeuProGlyLeuThrProSerAspAlaProLeuPro 69
Db 13849 GCGACGCGCCCGCGCAGCGCAAGCTCGGGTCGCCGCGCTGCCAGCGCCCGCGCCA 13790
QY 70 AlaLeu-ValIleHisGlyLeu-ThrPro-----ArgSerSerHisSerSera 85
Db 13789 GCAGCCACGCTCCACGGGTCTGCTCCACATCATCAAGCGCAGCGCGCTCGGGT 13730
QY 85 lacyLeuAlaSerAspSerGlyArgGluGlyGluArgGlyValArgGlyValArgThrHisC 105
Db 13729 GCTCGCTT-----CGCGCGCGCGCGCACAGCC 13703
QY 105 yshisargGlyIleGlyArgTrpValArgArgArgArgAsnGlyAlaAlaProGlyG 125
Db 13702 CCACAGCGC-----GCGCGCGCGAGGTCTCTGACGCGCTCTGTCGGG 13658
QY 125 luAlaProHisSerPro-----VallysluLysProValMetSerAsnIleGlyLys 143
Db 13657 CCGCGCCACCGCCCGCGGTGATCCACACAGCTCGTCTGAGCGCGCTCGG 13598
QY 143 erThrAsnIleLeuThrPhisAsnCysLeuIleGlylnSerAspArgGlnLysLeuG 163
Db 13597 ACAGC-----CACGCTGC-----AGCAGCGACAGCGCTCGTCTGTCG 13559
QY 163 lylGlnLysGlyCysValIleThrGlyLeuSerGlySerGlyLysSerThrLeuA 183
Db 13558 CCTCATGCGACGACCCCGCCACAGCTCCACGCGCTCGGCTCGCGCGTACGCTCGACCA 13499
QY 183 lacys-AlaLeuSerArgGluLeuHisCysArgGly-HisLeuThrTyrValLeuAsp 202
Db 13498 CCACCGCTCGAGCGCTCGCGCATGCTCGAGCGCGCACCAACCAATGCATCGAGATCG 13439
QY 202 yAspAsnLeuArg-----HisGlyLeuAsnArgAspLeuSerPheLysAlaGluAs 219
Db 13438 CAATCGCTCGCGCCCGCGGCTCGCGCGCGCGCTCC-----TCCCTTCGCTGCACCA 13382
QY 219 p-----ArgAlaGluAsnIleArgArgValGlyGluValAlaLysLeuPheAlaAs 236
Db 13381 TGACCAACGCGCGCGCTCACCAGGAGGTGCTCGAGGTCCACAGGTGCGAAGCTCA 13322
QY 236 pAlaGlyValIleCysIleAlaSerLeuIleSer-----ProTyr-----ArgAr 251
Db 13321 CCTG--ATCATATGTGGCCACAGCTGAACGCGGTCCGAGCTGCTCGCGCTCG 13265
QY 251 gAspArgAspAlaCysArgAlaLeuLeuProHisSerAsnPheIleGluValPheLeas 271
Db 13264 CCGCGCGCAGCTAGAGCGCGCATGCTCGCCAC-----CA 13229
QY 271 pLeuProLeuLysIleCysGluAlaArgAspProLysGlyLeuTyr----- 286
Db 13228 CCGCGCGCTGCACTGCTGACACGTACAGCAAGCGCGCTGCTGTCGCTGTTGTCCT 13169
QY 287 -----LysLeuAlaArgThrGlyLysIleLys 295
Db 13168 GCATCTCCACTCGGATCCGAAGCTCGTCCGCCCGCTCGCGCGCTGCGG 13123
RESULT 6
US-09-301-665-3
; Sequence 3, Application US/09301665
; Patent No. 620786
; GENERAL INFORMATION:
; APPLICANT: KELLEMS, RODNEY E.
; APPLICANT: DATTA, SURJIT K.
; APPLICANT: BLACKBURN, MICHAEL R.
; TITLE OF INVENTION: ADENOSINE DEAMINASE DEFICIENT TRANSGENIC MICE AND
```

```
; TITLE OF INVENTION: METHODS FOR THE USE THEREOF
; FILE REFERENCE: US/243
; CURRENT APPLICATION NUMBER: US/09/301.665
; EARLIER FILING DATE: 1999-04-28
; EARLIER APPLICATION NUMBER: 60/083,408
; EARLIER FILING DATE: 1998-04-29
; EARLIER APPLICATION NUMBER: 60/083,370
; EARLIER FILING DATE: 1998-04-28
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 36741
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-301-665-3

Alignment Scores: 16.6 Length: 36741
Score: 123.00 Matches: 64
Percent Similarity: 38.11% Conservative: 29
Best Local Similarity: 26.23% Mismatches: 70
Query Match: 6.74% Indels: 81
DB: 4 Gaps: 16

US-09-720-384a-4 (1-343) x US-09-301-665-3 (1-36741)
QY 11 ProLeuValThrHisThrGlnGlnProProSerPro---AlaProGlyProAlaSer--- 28
Db 25796 CCGCTGAGCTCCACCGG-----CCCTCTCCAGTCGCTGGGCTGGCAAGCCTC 25846
QY 29 -----GlnGlyGlnArgGlnGlyAsnThrLeuLeuSerProThrProThrLeuAlaVal 46
Db 25847 TGATGTCCAGCAGATGAGGGTCAAGTCTGATG---CCTGCCCTACCTCGGAATT 25903
QY 47 lIleLeuValAsnProGlnArgAlaProProValLeuLeuProGlyLeuThrProSerAspAla 66
Db 25904 GTGATGCTGAG-----TTACTGCCCTGATACCCCTGACTGGGCA 25945
QY 67 -----ProLeuPro-AlaLeu----- 71
Db 25946 TAGGACCAGCTGCTGAGCCAGCTCTGGGCTGAGGAGGAGCCATGACTTGACCTGG 26005
QY 72 -----ValII 73
Db 26006 CACTTTCCTTGTCTCCAGCATCAGTCAACCAAGATATGAGGGGGTGTGTGCATGTG 26065
QY 73 eHisGlyLeuThrProArgSerSerHisSerSerAlaGlyLeuAla-----Se 89
Db 26066 GCACACATACACACACACACACACACACACTTCAACCTGTTATCCCTTGAGATTTC 26125
QY 89 rAsp-----SerGlyArgArg-----GluGlyGluGlyArgGlyAlaArg---Th 103
Db 26126 TGACTTGTGCTATGGGGGTAGAAGGTGCTGGAAAAATTCGGTCTCTGTTCTCAGTTTC 26185
QY 103 rHisCysHisArgGlyIleGlyArgTrpValArgArgArgArg----- 117
Db 26186 CCATCTGTCAGTGGG---AGCAGCTGAGTAGAGAGAGCCCATGCTCTCTGCTGTGGTC 26242
QY 118 -----ArgAsnGlyAlaAlaProGlyGluAlaProHisSerProProLysGluLysProVa 136
Db 26243 CTGCAAGGAGGCTGGCGCTCTGCTGAGTCTGCTCCATCTGCGCTCTGAGCGCTGCTGAT 26302
QY 136 lMetSerAsnIleGlyLysSerThrAsnIleLeuThrPhisAsnCysLeuIleGlyGlnSe 156
Db 26303 CCGTCCCGGGTGGTCCACCACTCACTGTTTGTGTTTCCAGGA-----GGAGAGG 26353
QY 156 rAspArgGlnLysLeuLeuGlyGlnLysGlyCysValValIleThrThrGlyLeuSerG 176
Db 26354 ATCGCCCTCCAGCTAACACAGCAGAGGGGCTSC-----TGAACGCTATTGGCATG 26404
QY 176 ySerGlyLysSer-----ThrLeuAlaCysAlaLeuSe 187
Db 26405 GACAAGCCGCTCACCCCTCCAGACTTCTCGCCCAAGTTTGACTACTACTGCTGCTATC 26464
```


;;
;; CURRENT APPLICATION DATA:
;; FILING DATE: US/08/156,020
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Greenfield Ph.D., Michael S.
;; REGISTRATION NUMBER: 37,142
;; REFERENCE/DOCKET NUMBER: 93,413
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (312)715-1000
;; TELEFAX: (312)715-1234
;; INFORMATION FOR SEQ ID NO: 3:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 2626 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
;; ORGANISM: Thermus aquaticus
;; FEATURE:
;; NAME/KEY: mutation
;; LOCATION: replace(89, "g")
;; OTHER INFORMATION: /note= "This mutation results in a nucleotide
;; OTHER INFORMATION: alteration at position 89 of the native Tag DNA
;; OTHER INFORMATION: polymerase nucleotide sequence of C to G."
;; FEATURE:
;; NAME/KEY: mutation
;; LOCATION: replace(934, "a")
;; OTHER INFORMATION: /note= "This mutation results in a nucleotide
;; OTHER INFORMATION: alteration at position 934 of the native Tag DNA
;; OTHER INFORMATION: polymerase nucleotide sequence of T to A. This results in an
;; OTHER INFORMATION: amino acid change of Phe to Ile."
;; FEATURE:
;; NAME/KEY: mutation
;; LOCATION: replace(962, "c")
;; OTHER INFORMATION: /note= "This mutation results in a nucleotide
;; OTHER INFORMATION: alteration at position 962 of the native Tag DNA
;; OTHER INFORMATION: polymerase nucleotide sequence of T to C. This results in an
;; OTHER INFORMATION: amino acid change of Leu to Pro."
;; FEATURE:
;; NAME/KEY: mutation
;; LOCATION: replace(2535, "a")
;; OTHER INFORMATION: /note= "This mutation results in a nucleotide
;; OTHER INFORMATION: alteration at position 2535 of the native Tag DNA
;; OTHER INFORMATION: polymerase nucleotide sequence of G to A. This mutation is
;; OTHER INFORMATION: conservative."
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 121..2619
;; FEATURE:
;; NAME/KEY: mat_peptide
;; LOCATION: 121..2616
;; FEATURE:
;; NAME/KEY: -
;; LOCATION: 1..2619
;; OTHER INFORMATION: /note= "PLSM5"
;;

US-09-720-384A-4 (1-343) x US-08-156-020-3 (1-2626)

Alignment Scores:
Pred. No.: 0.538 Length: 2626
Score: 120.50 Matches: 85
Percent Similarity: 32.72% Conservative: 21
Best Local Similarity: 26.23% Mismatches: 128
Query Match: 6.60% Indels: 91
DB: 1 Gaps: 15

US-09-720-384A-4 (1-343) x US-08-156-020-3 (1-2626)

Qy 17 GlnGlnProSerProAlaProGlyProAlaSerGlnGlnArgGlnGlyAsnThr 36
||||: ||| ||||| ||||| ||| ||| |||

Db 1606 CAAAGAGGACCCCTTCCAGCT---GGTCCGGGAGTTGAGGTTGAAGGGGTGGCGGCCA 1550
Qy 37 LeuLeuSerProThrProThrLeuAlaValIleLeuValAsnProGlnArgAlaProPro 56
Db 1549 GCGGGAAGACCTCGGCCT-----CGAGCGGGGGAATCTCTCTCGG 1511
Qy 57 ValLeuProGlyLeuThrProSerAspAlaProLeuProAlaLeu----- 71
Db 1510 CCACCTCCAGGCAAGGCCCTTGAGATAGGCACGTCACAGGCGCACCCCGGTGGCTCCA 1451
Qy 72 -----ValIleHisGlyLeuThr 77
Db 1450 TGTGGCCAGCACAGCGAAAGGGCCCTCTCCACCTCCCGGTAAAGCCAAAGACCTCT 1391
Qy 78 ProArgSerSerHisSerSerAlaGly-----LeuAlaSerAspSerGlyArg----- 93
Db 1390 CCTCCCGCTCAAGCCTCCCCCACAGAGCCTCTCGGAAGAGCGGCCCGCT 1331
Qy 94 -----ArgGluGlyGluGlyArgGlyAlaAlaArgThrHisCys 105
Db 1330 CCCCCT 1271
Qy 106 HisArgGly-----IleGlyArgTrpValArgArgArg----- 117
Db 1270 TGGAGAGGTCCAGGAGTAGGCGAGGAGCATGGGTCTGCCGGCGGGAGGCCAAGGC 1211
Qy 118 -----ArgAsnGlyAlaAlaProGlyGluAla 126
Db 1210 CTTCCTCAGGCGCAGAGGCTCAGGTCTTTGGCGAAGAGCCCGCGCTCTCTTCAGGT 1151
Qy 127 Pro-----HisSerProValysGluLysProValMetSerAsnIle 140
Db 1150 CCCTGAGGGCTTTATAGGCTCGGGGGCCCGGTGACCGCCCGCCCTCTGGCGGGGCCA 1091
Qy 141 GlyLysSerThrAsnIleLeuTrpHisAsnCysLeuIleGlnSerAspArgGlnLys 160
Db 1090 GGGCCAGAAGATCGGCCACATGGGTCTCTTGGCGAAGAGCAAAAGCCAGGAAGGCC 1031
Qy 161 LeuLeuGlyGlnLysGlyCysValValTrpIleThrGlyLeuSerGlySerGlyLysSer 180
Db 1030 CTTC-GGC-----CGGGCCAGGGGCGCTCC 1005
Qy 181 ThrLeuAlaCysAlaLeuSerArgGluLeuHisCys---ArgGlyHisLeuThrTyrVal 199
Db 1004 TCCAGGGCTTGGGGCTTTCCAGAGGCGCAACTCGTGGAGGGG-----Leu 213
Qy 200 LeuAspGlyAspAsnLeuArgHis-----GlyLeuAsnArgAsp----- 213
Db 959 CTGCCAACTCAAGCTCTCCAGATGGCCCTTAAGCTCTCTCCGGTCGGGCTCCCGCTT 900
Qy 214 SerPhelysAlaGluAspArgAlaGluAsnIleArgValGlyGluValAlaLysLeu 233
Db 899 TTGGCGAAGTCCACCTCCAGGGGCGAGTCTGGTGGCACCTTTGGCCAGGTCCAGGAGC 840
Qy 234 PheAlaAspAlaGlyValIleCysIleAlaSerLeuIleSerProThrArgArgAspArg 253
Db 839 TTCAGATCGTCATGTGGCCAGGATCTTCCCGGATGGCGGTTCAGCGGTTCAGG 780
Qy 254 AspAlaCysArgAla---LeuLeuProHisSer-----AsnPhelGluValPheIle 270
Db 779 TTCTTGAGGAGGCTTCCAGGCTCCCGCTCCCGCTCCAGAGAGTCTCTCGCGCTCTCTCC 720
Qy 271 AspLeuProLeu 274
Db 719 CCGATGCCCTTG 708
RESULT 9
US-08-156-020-5/c
; Sequence 5, Application US/08156020
; Patent No. 5474920
; GENERAL INFORMATION:
; APPLICANT: Moses M.D., Robb E.
; TITLE OF INVENTION: Modified Thermo-Resistant DNA

```
; TITLE OF INVENTION: Polymerases
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allegretti & Witcoff
; STREET: 10 South Wacker Drive
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/156,020
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Greenfield Ph.D., Michael S.
; REGISTRATION NUMBER: 37,142
; REFERENCE/DOCKET NUMBER: 93,413
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312)715-1000
; TELEFAX: (312)715-1234
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2626 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Thermus aquaticus
; FEATURE:
; NAME/KEY: mutation
; LOCATION: replace(89, "g")
; OTHER INFORMATION: /note="This mutation results in a nucleotide
; OTHER INFORMATION: alteration at position 89 of the native Tag DNA
; OTHER INFORMATION: polymerase nucleotide sequence of C to G."
; FEATURE:
; NAME/KEY: mutation
; LOCATION: replace(934, "a")
; OTHER INFORMATION: /note="This mutation results in a nucleotide
; OTHER INFORMATION: alteration at position 934 of the native Tag DNA
; OTHER INFORMATION: polymerase nucleotide sequence of T to A. This results in an
; OTHER INFORMATION: amino acid change of Phe to Ile."
; FEATURE:
; NAME/KEY: mutation
; LOCATION: replace(962, "c")
; OTHER INFORMATION: /note="This mutation results in a nucleotide
; OTHER INFORMATION: alteration at position 962 of the native Tag DNA
; OTHER INFORMATION: polymerase nucleotide sequence of T to C. This results in an
; OTHER INFORMATION: amino acid change of Leu to Pro."
; FEATURE:
; NAME/KEY: mutation
; LOCATION: replace(2535, "a")
; OTHER INFORMATION: /note="This mutation results in a nucleotide
; OTHER INFORMATION: alteration at position 2535 of the native Tag DNA
; OTHER INFORMATION: polymerase nucleotide sequence of G to A. This mutation is
; OTHER INFORMATION: conservative."
; FEATURE:
; NAME/KEY: mutation
; LOCATION: replace(193, "t")
; OTHER INFORMATION: /note="This mutation changes the nucleotide
; OTHER INFORMATION: alteration at position 193 of the native Tag DNA polymerase from
; OTHER INFORMATION: C to T, resulting in an amino acid change of Arg to Cys."
; FEATURE:
; NAME/KEY: mutation
; LOCATION: replace(504, "a")
; OTHER INFORMATION: /note="This mutation changes the nucleotide

; OTHER INFORMATION: at position 504 of the native Tag DNA polymerase from
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 121..2619
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 121..2616
; FEATURE:
; NAME/KEY:
; LOCATION: 1..2619
; OTHER INFORMATION: /note="ptarf3"
US-08-156-020-5

Alignment Scores:
Pred. No.: 0.538 Length: 2626
Score: 120.50 Matches: 85
Percent Similarity: 32.72% Conservative: 21
Best Local Similarity: 26.23% Mismatches: 128
Query Match: 6.60% Indels: 91
DB: 1 Gaps: 15

US-09-720-384A-4 (1-343) x US-08-156-020-5 (1-2626)
QY 17 GlnGlnProSerProAlaProGlyProAlaSerGlnGlnGlyAlaArgGlnGlyAsnThr 36
|||::: ||| ||||| ||||| ||| ||| ||| |||
DB 1606 CAAAGAGGACCCCTTCCAGCT---GGTCCGGGAGTTGAGGTTGAAGGGGTGGCGCGCCA 1550
QY 37 LeuLeuSerProThrProThrLeuAlaValIleLeuValAsnProGlnArgAlaPro 56
||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1549 GCGGAGACCTCGGCCT-----CGAGGGGGGGATCTCTCCGCG 1511
QY 57 ValLeuProGlyLeuThrProSerAspAlaProLeuProAlaLeu----- 71
||| ||| ||| ||| ||| ||| ||| |||
DB 1510 CCACCTCCAGGACAAAGGCCCTTGATAGGCCAGCTCCAGGCGCACCCCGCTGGCGCTCCA 1451
QY 72 -----ValIleHisGlyLeuThr 77
||| ||| ||| |||
DB 1450 TGTGGGCCAGGACAGCGGAAAGGGGCCTCTCCACCTCCCGGTAAAGCCAAAGGAGCCTCT 1391
QY 78 ProArgSerSerHisSerSerAlaGly-----LeuAlaSerAspSerGlyArg----- 93
||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1390 CCTCCCGCTCAAGCCTCCCCCACAGGTTGGCGAAGAGCCTCTCGGAAAGGGGGCGCGCT 1331
QY 94 -----ArgGluGlyGluGlyArgGlyAlaArgThrHisCys 105
||| ||| ||| ||| ||| ||| ||| |||
DB 1330 CCCCCTCTCTCCGTCCTCCGTCCTCCCGCGGTAGCGCGGCGCACCCCTCGGGGGTGTGT 1271
QY 106 HisArgGly-----IleGlyArgTrpValArgArgArg----- 117
::: ||| ||| ||| ||| ||| ||| ||| |||
DB 1270 TGGAAAGGTCCTCAGGAGGTAGCGGAGGAGCATGTGGGTCTGTGCGCGGGCGGAGGCCAAGGC 1211
QY 118 -----ArgAsnGlyAlaAlaProGlyGluAla 126
||| ||| ||| ||| ||| ||| ||| |||
DB 1210 CTTCCCTCAGGGCCACAGCGCTCAGGCTTTTGGCGGAGAGAGCCCGCGCGCTCTTCAGGT 1151
QY 127 Pro-----HisSerProValLysGluLysProValMetSerAsnIle 140
||| ||| ||| ||| ||| ||| ||| |||
DB 1150 CCCTGAGGGCTTTAATAGSCTCGGGGGCCCGGTGGACCCCGCGCCCTCGCGCGGCGCCA 1091
QY 141 GlyLysSerThrAsnIleLeuThrHisAsnCysLeuIleGlyGlnSerAspArgGlnLys 160
||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1090 GGGCCAGAAGTCGGCCACATGGGCTCTTGGCGGAAAGCACAAAGCCCAAGGCGCC 1031
QY 161 LeuLeuGlyGlnLysGlyCysValValTrpIleThrGlyLeuSerGlySerGlyLysSer 180
||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1030 CTTCC-GGC-----GGGGGCCAGGGGCGCTCC 1005
QY 181 ThrLeuAlaCysAlaLeuSerArgGluLeuHisCys---ArgGlyHisLeuThrTyrVal 199
::: ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1004 TCCAGGCGCTTGGGCTTTCCAGAGGCCCGCAACTCGTGGAGGGG----- 960
QY 200 LeuAspGlyAspAsnLeuArgHis-----GlyLeuAsnArgAsp-----Leu 213
```

```

Db 959 CTGCCAAACTAAGCCTCTCCAGAAATGGCCCTAAGCCTCTCCCGGTCCGGCTCCCGCCTT 900
Qy 214 SerPhelyAlaGluAspArgValGluValAlaLysLeu 233
Db 899 TTGGCAAGTCCACCTCCAGGGCAGGTGCGTGGCAGCTTGGCAGTCCAGGAGC 840
Qy 234 PheAlaAspAlaGlyValIleCysIleAlaSerLeuIleSerProTyrArgAspArg 253
Db 839 TTGATGTCGTGGCCAGGATCTTCCCGGATGGCGGCTTCAGCCGTCACG 780
Qy 254 AspAlaCysArgAla---LeuLeuProHisSer-----AspPheIleGluValPheIle 270
Db 779 TTTCTGAGGAGGCTTCCAGGCTCCCACTCTCTCCAGAAGCTTCTTCGCGCTTCTCC 720
Qy 271 AspLeuProLeu 274
Db 719 CCGATGCCCTTG 708

RESULT 10
US-08-156-020-7/c
; Sequence 7, Application US/08156020
; Patent No. 5474920
; GENERAL INFORMATION:
; APPLICANT: Moses M.D., Robb E.
; TITLE OF INVENTION: Modified Thermo-Resistant DNA
; TITLE OF INVENTION: Polymerases
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allegretti & Witcoff
; STREET: 10 South Wacker Drive
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/156,020
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Greenfield Ph.D., Michael S.
; REGISTRATION NUMBER: 37,142
; REFERENCE/DOCKET NUMBER: 93,413
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312)715-1000
; TELEFAX: (312)715-1234
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2626 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORGANISM: Thermus aquaticus
; ORGANISM: Thermus aquaticus
; FEATURE:
; NAME/KEY: mutation
; LOCATION: replace(89, "g")
; OTHER INFORMATION: /note= "This mutation results in a nucleotide
; OTHER INFORMATION: alteration at position 89 of the native Tag DNA
; OTHER INFORMATION: polymerase nucleotide sequence of C to G."
; FEATURE:
; NAME/KEY: mutation
; LOCATION: replace(934, "a")
; OTHER INFORMATION: /note= "This mutation results in a nucleotide
; OTHER INFORMATION: alteration at position 934 of the native Tag DNA

```

```

; OTHER INFORMATION: polymerase nucleotide sequence of T to A. This results in
; OTHER INFORMATION: amino acid change of Phe to Ile."
; FEATURE:
; NAME/KEY: mutation
; LOCATION: replace(962, "c")
; OTHER INFORMATION: /note= "This mutation results in a nucleotide
; OTHER INFORMATION: alteration at position 962 of the native Tag DNA
; OTHER INFORMATION: polymerase nucleotide sequence of T to C. This results in
; OTHER INFORMATION: amino acid change of Leu to Pro."
; FEATURE:
; NAME/KEY: mutation
; LOCATION: replace(2535, "a")
; OTHER INFORMATION: /note= "This mutation results in a nucleotide
; OTHER INFORMATION: alteration at position 2535 of the native Tag DNA
; OTHER INFORMATION: polymerase nucleotide sequence of G to A. This mutation i
; OTHER INFORMATION: conservative."
; FEATURE:
; NAME/KEY: mutation
; LOCATION: replace(341, "a")
; OTHER INFORMATION: /note= "This mutation results in a nucleotide
; OTHER INFORMATION: alteration at position 341 of the native Tag DNA
; OTHER INFORMATION: polymerase nucleotide sequence of G to A. This mutation
; OTHER INFORMATION: results in an amino acid change of Arg to His."
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 121..2619
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 121..2616
; FEATURE:
; NAME/KEY: -
; LOCATION: 1..2619
; OTHER INFORMATION: /note= "pTarf5"
; US-08-156-020-7

Alignment Scores:
Pred. No.: 0.538 Length: 2626
Score: 120.50 Matches: 85
Percent Similarity: 32.72% Conservative: 21
Best Local Similarity: 26.23% Mismatches: 128
Query Match: 6.60% Indels: 91
Db: 1 Gaps: 15

US-09-720-384A-4 (1-343) x US-08-156-020-7 (1-2626)
Qy 17 GlnGlnProSerProAlaProGlyProAlaSerGlnGlyGlnArgGlnGlyAsnThr 36
Db 1606 CAAAGAGGACCCCTTCCAGCT---GGTCCGGGAGTTGAGTTGAAGGGTGGCCGGCCA 1550
Qy 37 LeuLeuSerProThrProThrLeuAlaValIleLeuValAsnProGlnArgAlaProPro 56
Db 1549 GCGGAAGACCTCGGCCT-----CGAGGGCGGCGATCTCTCTCGG 1511
Qy 57 ValLeuProGlyLeuThrProSerAspAlaProLeuProAlaLeu----- 71
Db 1510 CCACCTCCAGGGACAAGCCCTGAGATAGGACGCTCCAGGCGCACCCCGCTGGCCTCCA 1451
Qy 72 -----ValIleHisGlyLeuThr 77
Db 1450 TGTGGGCCAGGACAGCGGAAAGGGCCCTCTCCACCTCCCGGTAAAGCCAAAGGAGCCTCT 1391
Qy 78 ProArgSerSerHisSerSerAlaGly-----LeuAlaSerAspSerGlyArg----- 93
Db 1390 CCTCCCTCCAGCCCTCCCGCCACAGCTTGGCGAAGAGACCTCTCGGAAAGGCGCGCCGCT 1331
Qy 94 -----ArgGluGlyGluGlyArgGlyAlaArgThrHisCys 105
Db 1330 CCCCCTCTCTCCGCTCCACTCCCGCGGTAGCGCGGCGCACCCCTCTCGGGGTGTGT 1271
Qy 106 HisArgGly-----IleGlyArgTrpValArgArgArg----- 117
Db 1270 TGGAAAGGTCAGGAGGTAGGCGGAGGACGTGGGGTCTGTCGCGGGCGGAGGCAAGGC 1211

```



```
Db 1510 CCACCTCCAGGACAGGCCCTGAGATAGCCACGCTCCAGGCGCACCCCGCTGCCTCCA 1451
QY 72 -----ValleHisGlyLeuThr 77
Db 1450 TGTGGCCAGGACAGCGAAAGGGCCCTCTCCACCTCCGCGTAAAGCAAGAGCGCTCT 1391
QY 78 ProArgSerSerHisSerAlaGly-----LeuAlaSerAspSerGlyArg----- 93
Db 1390 CTTCCCTCAAGCCCTCCCCACAGGTTGGCGAAGAGCTCTCGGAAGAGGCGCGCGCT 1331
QY 94 -----ArgGluGlyGluGlyArgGlyAlaAArgThrHisCys 105
Db 1330 CCCCCTCCTCCTCGCTCCACTCCCGCGCTAGCGCGCGCCACCCCTCGGGGTGTGT 1271
QY 106 HisArgGly-----IleGlyArgTrpValArgArgArg----- 117
Db 1270 TGAAGGTTCCAGGAGGTAGCGAGAGCATGGGGTCTCGCGCGCGGAGGCGCAAGGC 1211
QY 118 -----ArgAsnGlyAlaAlaProGlyGluAla 126
Db 1210 CTTCCCTCAGGCGCAGACGCTCAGGTCTTTGGCGAGAGCCCGCGCTCTCTCAGGT 1151
QY 127 Pro-----HisSerProValGlySerProValMetSerAsnIle 140
Db 1150 CCTGAGGCTTTAAGGCTCGGGGCGCGGTGGACCGCGCCCTCGCGCGCGGCCA 1091
QY 141 GlyLysSerThrAsnIleLeuTrpHisAsnGlyLeuIleGlyGlnSerAspArgGlnLys 160
Db 1090 GGGCCAGACAGTGGGCCACATGGGCTCTTGGGGAAAGCACAAAGCCACGAAGGCC 1031
QY 161 LeuLeuGlyGlnLysGlyCysValValTrpIleThrGlyLeuSerGlySerGlyLysSer 180
Db 1030 CTTC-GGC-----GGGGCGAGGGGGCGCTCC 1005
QY 181 ThrLeuAlaCysAlaLeuSerArgGluLeuHisCys---ArgGlyHisLeuThrTyrVal 199
Db 1004 TCCAGGCGCTGGGGCTTTCCAGAGGCGCAACTCGTGGAGGGG----- 960
QY 200 LeuAspGlyAspAsnLeuArgHis-----GlyLeuAsnArgAsp-----Leu 213
Db 959 GTGCAAACTCAAGCTCTCCAGATGGCTTAAGCTCTCCCGGTGGGCTCCCGCTT 900
QY 214 SerPheLysAlaGluAspArgAlaGluAsnIleArgArgValGlyGluValAlaLysLeu 233
Db 899 TTGCGAAGTCCACCTCCAGGGGCGAGTCTGCGGACCTTGGCCAGGTCCCGAGGAGC 840
QY 234 PheAlaAspAlaGlyValIleCysIleAlaSerLeuIleSerProTyrArgArgAspArg 253
Db 839 TTCAGATCGTCCATGTGGCCAGGATCTTCTCCGGATGGCGGCTTCAGCGGTCCAGG 780
QY 254 AspAlaCysArgAla---LeuLeuProHisSer-----AsnPheIleGluValPheIle 270
Db 779 TTTGTAGGAGGCTTCCAGGCTCCCGCTCCCGCTCCCGAGAGCTCTCTCGCGCTTCTCC 720
QY 271 AspLeuProLeu 274
Db 719 CCGATGCCCTTG 708
RESULT 12
US-08-923-137-2/c
; Sequence 2, Application US/08923137
; Patent No. 6083716
; GENERAL INFORMATION:
; APPLICANT: Willson, James M.
; APPLICANT: Fatina, Steven F.
; APPLICANT: Fisher, Krishna J.
; TITLE OF INVENTION: Chimpanzee Adenovirus Vectors
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Cntr., P.O. Box 457
; CITY: Spring House
; STATE: Pennsylvania
```

```
; COUNTRY: United States of America
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/923.137
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/024,700
; FILING DATE: 06-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: GNPVN.021CIPUSA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9200
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 36519 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; US-08-923-137-2

Alignment Scores:
Pred. No.: 32.9 Length: 36519
Score: 119.50 Matches: 70
Percent Similarity: 34.96% Conservative: 23
Best Local Similarity: 26.32% Mismatches: 90
Query Match: 6.54% Indels: 84
DB: Gaps: 12

US-09-720-384A-4 (1-343) x US-08-923-137-2 (1-36519)
QY 19 ProProSerProAlaProGlyProAlaSerGlnGlnArgGlnGlyAsnThrLeuLeu 38
Db 15881 CCTCGTCCCGTGGCTGGCTCGCGCAGCAGGCTCGCGCGCGCGCGCGCGGGT 15822
QY 39 SerProThrProThrLeuAlaValIleLeuValAsnProGlnArgAlaProValLeu 58
Db 15821 GCTCCGCTGGCG----- 15810
QY 59 ProGlyLeuThrProSerAspAlaProLeuProAlaLeuValIleHisGlyLeuThrPro 78
Db 15809 -----CGGGCGGATGGCGCGCGCGCTCTTGGCGCG---GGCGTACCGGCG 15765
QY 79 ArgSerSerHisSerSerAlaGlyLeuAlaSerAspSerGlyArgArgGluGlyGluGly 98
Db 15764 CGCGTCCCGCACAC-----GCTGTGATGACGCGCTCCACCGTGAGACGGG 15717
QY 99 ArgGly-----AlaArgThrHisCysHisArgGlyIleGlyArgTrpVal 113
Db 15716 CGCGCGCGGGGGTGTAGTTGGCGCGCTCGCGCACACCT-GGT----- 15673
QY 114 ArgArgArgArgArgAsnGlyAlaAlaProGlyGluAlaProHisSerProValLysGlu 133
Db 15672 CGATCAGCTGTGTGACGCTGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGGAG 15613
QY 134 LysProValMetSerAsnIleGlyLysSerThrAsnIleLeuTrpHisAsnCysLeuIle 153
Db 15612 CGCGGA-----AGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 15568
QY 154 GlyGlnSerAspArgGln----- 159
Db 15567 GCGAGAGCGCTTCCGTACATCTTGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 15508
QY 160 -----LysLeuLeuGlyGlnLysGly----- 166
```

Db 15507 GCGAGATGAGATGACATTAGAGGTCGGCTCGAGAGGACGCGCGCGACTATGC 15448
Qy 167 -----CysValValTrpIleThrGlyLeuSerGlySer 177
Db 15447 CCAGGGCCCTTGTAGACGTAGGGGAGGTGGCGGCTCTGGCTCAGTAACGGTCACGCGCT 15388
Qy 178 GlySerThrLeuAlaCysAlaLeu-----SerArgGluLeu-HisCysArgGlyH1 195
Db 15387 GGACTCCCGGACTACTGTCGCGCAGCGGAGGTCCCGTGATCTGTGAGACGAGAACCT 15328
Qy 195 sLeuThrTyrrValLeuaspGlyaspAsnLeuArgHisGlyLeuAsnArgaspLeuSerPh 215
Db 15327 TTTCACCT-----GACGGTGGTAATGGTGGCGCGCGG---CGGCGGACGAGGATCT 15280
Qy 215 eLysAlaGluAspArgAlaGluAsnIleArgArgValGlyGluValAlaIstLeuPheAl 235
Db 15279 GGTTCCTCGGGGAAGCGGTTGAAGACGTGCGTAAGCGAGGTGAAGCGCGCGAGCTGCTGCG 15220
Qy 235 aAspAlaGlyValIle 240
Db 15219 AGTAGACGCGCTGCTC 15204

RESULT 13
US-07-602-848E-1/C
; Sequence 1, Application US/07602848E
; Patent No. 6083686
; GENERAL INFORMATION:
; APPLICANT: Sullivan, Mark Alan
; TITLE OF INVENTION: INCREASED PRODUCTION OF
; TITLE OF INVENTION: THERMUS AQUATICUS DNA POLYMERASE IN E. COLI
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eastman Kodak Company,
; ADDRESS: Patent Department
; STREET: 343 State Street
; CITY: Rochester
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 1 4 6 5 0 - 2 0 1
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch,
; MEDIUM TYPE: 1.44 MB storage (IBM)
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS DOS Version 3.3
; SOFTWARE: PC-8 (Word for Windows)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/602,848E
; FILING DATE: 19901026
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA: No. 6083686
; ATTORNEY/AGENT INFORMATION:
; NAME: Tucker, J. Lanny
; REGISTRATION NUMBER: 27,678
; REFERENCE/DOCKET NUMBER: 58374
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 722-9332
; TELEFAX: (716) 477-4646
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2499
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: genomic DNA
; HYPOTHETICAL: no
; ANTI-SENSE: no
; ORIGINAL SOURCE:
; ORGANISM: Thermus aquaticus
; INDIVIDUAL ISOLATE: Yrl, ATCC 25104
; IMMEDIATE SOURCE: amplified from genomic DNA
; FEATURE:
; NAME/KEY: peptide

; LOCATION: 1-2496
; IDENTIFICATION METHOD: comparison to
; IDENTIFICATION METHOD: sequence in GenBank, Accession number J04639.
; PUBLICATION INFORMATION:
; AUTHORS: Lawyer, F.C., Stoffel, S.,
; AUTHORS: Saiki, R.K., Myambo, K., Drummond, R.,
; AUTHORS: Gelfand, D.H.
; TITLE: Isolation, characterization and
; TITLE: expression in Escherichia coli of the DNA
; TITLE: polymerase gene from Thermus aquaticus.
; JOURNAL: Journal of Biological
; JOURNAL: Chemistry
; VOLUME: 264
; ISSUE: 11
; PAGES: 6427-6437
; DATE: 15 April 1989
US-07-602-848E-1
Alignment Scores:
Pred. No.: 0.673 Length: 2499
Score: 119.00 Matches: 83
Percent Similarity: 31.12% Conservative: 20
Best Local Similarity: 25.08% Mismatches: 124
Query Match: 6.52% Indels: 105
DB: 3 Gaps: 13
US-09-720-384A-4 (1-343) x US-07-602-848E-1 (1-2499)
Qy 17 GlnGlnProProSerProAlaProGlyProAlaSerGlnGlyGlnArgGlnGlyAsnThr 36
Db 1486 CAAAGAGGACCTTTCACGCT---GGTCCCGGGAGTTGAGGTGAGGGGGTGGCGGCCA 1430
Qy 37 LeuLeuSerProThrProThrLeuAlaValIleLeuValAsnProGlnArgAlaProPro 56
Db 1429 GCGGAAGACCTCGGCT-----CGAGGGCGGCGATCTCTCTCGG 1391
Qy 57 ValLeuProGlyLeuThrProSerAspAlaProLeuProAlaLeu----- 71
Db 1390 CCACCTCCAGGGACAGAGGCTTGAGTAGGCGACGTCAGGCGCCACCCCGCTGGCGTCCA 1331
Qy 72 -----ValIleHisGlyLeuThr 77
Db 1330 TGTGGCGCAGGACAGCGAAAGGGGCTCTCCACCTCCCGGTAAAGCCAAAGAGGCTCT 1271
Qy 78 ProArgSerSerHisSerSerAlaGly-----LeuAlaSerAspSerGlyArg----- 93
Db 1270 CTTCCCTCAAGCTTCCCTCCACACAGGTTGGCGAAGACCTCTCGGAAAGGGCGCGCT 1211
Qy 94 -----ArgGluGlyGluGlyArgGlyAlaArgThrHisCys 105
Db 1210 CCCCCTCTCTCTCCCTCCACTCCCGCGCTAGCGCGGCGCCACCCCTCGGGGGTGTGT 1151
Qy 106 HisArgGly-----IleGlyArgTrpValArgArgArg----- 117
Db 1150 TGGAGGGTCCAGGAGGTAGCGAGGACGATGGGTCTCGCGGGGGAGGCGCAAGGC 1091
Qy 118 -----ArgAsnGlyAlaAlaProGlyGluAla 126
Db 1090 CTTCCCTCAGGCGCAGACGCTCAGGTCTTTGGCGAGAAAGCCCGCGCTCTCTCAGGT 1031
Qy 127 Pro-----HisSerProValLysGluLysProValMetSerAsnIle 140
Db 1030 CCCTGAGGGCTTTATAAGGCTCGGGGCGCGGTGGAGCCCGCGCCCTCGGGCGGCGCA 971
Qy 141 GlyLysSerThrAsnIleLeuTrpHisAsnCysLeuIleGlyLysSerAspArgGlnLys 160
Db 970 GGGCGAGGAGATCGGCGCCACATGGGCTCTTGGCGGGAAGACACAAAGCCCGAGGCGCC 911
Qy 161 LeuLeuGlyGlnLysGlyCysValValTrpIleThrGlyLeuSerGlySerGlyLysSer 180
Db 910 CTTCC-GGC-----GGGGCGCGAGGGGCGCTCC 885
Qy 181 ThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTyrrValLeu 200


```

      :: ||| |||||
Db 884 TCAGGCGCTTGGGCTTCCAGAAGG----- 858
QY 201 AspGlyAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLysAla----- 217
      ::::: ||| ||||| |||||
Db 857 ---CCGACTCGTGGAGGAGCTGCCAACTCAAGCTCTCCAGAAAGCCCTAAGCCCC 801
QY 218 -----GluAspArgAlaGluAsnIleArgArg 226
      ::::: ||| |||||
Db 800 TCCGGCTCGGGCTCCCGCTTTTGGCGAAGTCCACCTCCAGGGCAGGTCCGTGCGCACC 741
QY 227 ValGlyGluValAlaLysLeuPheAlaAspAlaGlyValIleCysIleAlaSerLeuIle 246
      ::::: ||| |||||
Db 740 TTGGCCAGGTCCAGGAGCTTCAGATCGTCCATCTGGGCCAGGATCTTCCCGCATG 681
QY 247 SerProTyrArgArgAspAlaCysArgAla---LeuLeuProHisSer----- 263
      ::::: ||| |||||
Db 680 GCGGGCTTCAGCGGTCCAGGTCTTTCAGGAGGGCTTCCAGGCTCCCCCACTCTCCAGA 621
QY 264 AsnPheIleGluValPheIleAspLeuProLeu 274
      ::|||::: ||||| ::|||:::
Db 620 AGCTTCCTCGCGCTCTTCTCCCGCATGCCCTTG 588

```

RESULT 14

```

US-08-073-384C-9/c
; Sequence 9, Application US/08073384C
; Patent No. 5541311
; GENERAL INFORMATION:
; APPLICANT: Dahlberg, James E.
; APPLICANT: Lyamichev, Victor I.
; APPLICANT: Brow, Mary Ann D.
; TITLE OF INVENTION: SYNTHESIS-DEFICIENT THERMOSTABLE DNA
; TITLE OF INVENTION: POLYMERASE
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HAVERTOCK, MEDLEN & CARROLL
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/073,384C
; FILING DATE: 04-JUN-1993
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/986,330
; FILING DATE: 07-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Peter G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: FORS-00613
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/705-8410
; TELEFAX: 415/397-8338
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1647 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-073-384C-9
Alignment Scores:
Pred. No.: 0.538
Score: 117.00
Percent Similarity: 32.51%
Length: 1647
Matches: 84
Conservative: 21

```

```

Best Local Similarity: 26.01% Mismatches: 130
Query Match: 6.41% Indels: 89
DB: 1 Gaps: 14
US-09-720-384A-4 (1-343) x US-08-073-384C-9 (1-1647)
QY 17 GlnGlnProSerProAlaProGlyProAlaSerGlnGlyGlnArgGlnGlyAsnThr 36
      |||::: ||| ||||| ||||| ||| ||| |||
Db 1489 CAAAGAGAGACCTTTCCAGCT---GGTCCGGGAGTTGAGGTTGAAGGGTGGCGGCCA 1433
QY 37 LeuLeuSerProThrProThrLeuAlaValIleLeuValAsnProGlnArgAlaPro 56
      ||| ||| ||| |||::: |||
Db 1432 GCGGAAGACCTCGGCT-----CGAGCGGGGATCTCCCGG 1394
QY 57 ValLeuProGlyLeuThrProSerAspAlaProLeuProAlaLeu----- 71
      ||||| ||| ||| ||| |||
Db 1393 CCACCTCCAGGGACAAGGCCCTGAGATAGGCACGTCACGGCGCACCCCGCTGGCTCCA 1334
QY 72 -----ValIleHisGlyLeuThr 77
      ||| ||| ::|
Db 1333 TGTGGCCAGGACACGGAAGGGCCCTCTCCACCTCCCGTAAACCAAGAGACCTCT 1274
QY 78 ProArgSerSerHisSerSerAlaGly-----LeuAlaSerAspSerGlyArg----- 93
      ||| ||| ||| ||||| |||||
Db 1273 CTTCCCTCAAGCCTCCCCCACAGGTTGGGAAGAGACCTCTCGAAAGGGCGGCCGCT 1214
QY 94 -----ArgGluGlyGluGlyAlaArgThrHisCys 105
      ||| ||| ||| |||
Db 1213 CCCCCTCTCTCCGCTCCTCCCGCTAGCGCGGCCACCCCTCGGGGTGGTGT 1154
QY 106 HisArgGly-----lleGlyArgTrpValArgArgArg----- 117
      ::|||:: ||| ||| ||||| |||||
Db 1153 TCGAAGGTCCAGAGTAGGCGAGGAGCATGGGGTCTGCGCGGCGGAGGCCAAGGCC 1094
QY 118 -----ArgAsnGlyAlaAlaProGlyGluAla 126
      ||| ||| |||||
Db 1093 CTTCCCTCAGGGCCAGACGCTCAGGTCTTTGGCGAAGACCCCGCGCTCTTCAGGT 1034
QY 127 Pro-----HisSerProValLysGluLysProValMetSerAsnIle 140
      ||| ||| |||
Db 1033 CCCTGAGGGCTTTATAGGCTCGGGGGCCCGGTGGACCCCGCCCTTGGCGGGGCCA 974
QY 141 GlyLysSerThrAsnIleLeuThrHisAsnCysLeuIleGlnSerAspArgGlnLys 160
      ||| ||| ||| ||||| |||||
Db 973 GGGCCAGAAGATCGGCCCATATGGCTCTCTTCGGGAAGACACAAAGCCCAAGAGCCC 914
QY 161 LeuLeuGlnLysGlyCysValValTrpIleThrGlyLeuSerGlySerGlyLysSer 180
      ||| ||| ||| ||| ||| |||
Db 913 CTTCC-GGC-----GGGGCCAGGGGCGCTCC 888
QY 181 ThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTyrValLeu 200
      ::| ||| ||||| ||||| ||| ::|
Db 887 TCCAGGGCTTGGGGCTTTCCAGA-----AGGCGCAACTCGTGGAGGAGGTG 840
QY 201 AspGlyAspAsnLeuArgHis-----GlyLeuAsnArgAsp-----LeuSer 214
      ::|||:: ||| ||||| |||||
Db 839 CCAACTCAAGCTCTCCAGAAAGCCCTACCTCTCCCGGTGGGCTCCCGCTTTTG 780
QY 215 PheLysAlaGluAspArgAlaGluAsnIleArgValGlyGluValAlaLysLeuPhe 234
      |||::: ||| ::|||:: ||| ::|
Db 779 GCGAAGTCCACCTCCAGGGCAGGTGCGTGGCCACCTTGGCCAGGTCCCGAGGAGCTTC 720
QY 235 AlaAspAlaGlyValIleCysIleAlaSerLeuIleSerProTyrArgArgAspAsp 254
      ::| ||| ||| ||| ||| ||| |||
Db 719 AGATCGTCCATGTGGGCCAGGATCTTCCCGGATGGCGGCTTACCGGTCCAGGTTC 660
QY 255 AlaCysArgAla---LeuLeuProHisSer-----AsnPheIleGluValPheIleAsp 271
      ||||| ||||| ||||| |||||
Db 659 TTGAGGAGGGCTTCCAGGCTCCCCCACTCTCTCCAGAAAGCTCTCTCGCGCTTCTTCCCGG 600
QY 272 LeuProLeu 274
      ::|||::: |||
Db 599 ATGCCCTTG 591

```

RESULT 15
 US-08-254-359A-9/c
 ; Sequence 9, Application US/08254359A
 ; Patent No. 5614402
 ; GENERAL INFORMATION:
 ; APPLICANT: DAHLBERG, JAMES E.
 ; APPLICANT: LYAMICHEV, VICTOR I.
 ; APPLICANT: BROW, MARY ANN D.
 ; TITLE OF INVENTION: 5' NUCLEASES DERIVED FROM THERMOSTABLE
 ; TITLE OF INVENTION: DNA POLYMERASE
 ; NUMBER OF SEQUENCES: 40
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: HAVERSTOCK, MEDLEN & CARROLL
 ; STREET: 220 MONTGOMERY STREET, SUITE 2200
 ; CITY: SAN FRANCISCO
 ; STATE: CALIFORNIA
 ; COUNTRY: UNITED STATES OF AMERICA
 ; ZIP: 94104
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/254,359A
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/073,384
 ; FILING DATE: 06-JUN-1993
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/986,330
 ; FILING DATE: 07-DEC-1992
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: CARROLL, PETER G.
 ; REGISTRATION NUMBER: 32,837
 ; REFERENCE/DOCKET NUMBER: FORS-01000
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (415) 705-8410
 ; TELEFAX: (415) 397-8338
 ; INFORMATION FOR SEQ ID NO: 9:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1647 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; US-08-254-359A-9
 Alignment Scores:
 Pred. No.: 0.538 Length: 1647
 Score: 117.00 Matches: 84
 Percent Similarity: 32.51% Conservative: 21
 Best Local Similarity: 26.01% Mismatches: 130
 Query Match: 6.41% Indels: 89
 DB: 1 Gaps: 14
 US-09-720-384A-4 (1-343) x US-08-254-359A-9 (1-1647)
 Qy 17 GlnGlnProSerProAlaProGlyProAlaSerGlnGlnArgGlnGlyAsnThr 36
 Db 1489 CAAAGAGGACCTTTCCAGCT---GGTCCGGGAGTTGAGGTTGAAGGGGTGCCGGCCA 1433
 Qy 37 LeuLeuSerProThrLeuAlaValIleLeuValAsnProGlnArgAlaProPro 56
 Db 1432 GCGGGAAGACCTTCGGCCT-----CGAGCGGGCGATCTCCCCCG 1394
 Qy 57 ValLeuProGlyLeuThrProSerAspAlaProLeuProAlaLeu----- 71
 Db 1393 CCACCTCCAGGACAAAGCCCTGAGATAGGCCACCGTCCAGCGCACCCCGTGGCCTCCA 1334
 Qy 72 -----ValIleHisGlyLeuThr 77

Db 1333 TGTGGCCAGGACAGCGAAAGGGCCCTCTCCACCTCCCGGTAAGCCAAAGAGGAGCCTCT 1274
 Qy 78 ProArgSerSerHisSerSerAlaGly-----LeuAlaSerAspSerGlyArg----- 93
 Db 1273 CCTCCCTCAAGCCTCCCCACAGGTTGGCGAAGAGCCTCTCGGAAGGGCGGCCGCT 1214
 Qy 94 -----ArgGluGluGluGlyArgGlyAlaArgThrHisCys 105
 Db 1213 CCCCGGCTCTCCGTCCACTCCCGCGTAGCGCGGCCGCCACCCCTCGGGGTGGTGT 1154
 Qy 106 HisArgGly-----IleGlyArgTrrpValArgArgArg----- 117
 Db 1153 TGGAAGGTCCAGGAGGTAGCGAGGACATGGGTCTGCGCGGGGAGGCCCAAGGC 1094
 Qy 118 -----ArgAsnGlyAlaAlaProGlyGluAla 126
 Db 1093 CTTCCCTCAGGGCCAGACGCTAGGTCTTGGCGAGAGCCCGCGCCTCTCTTCAGGT 1034
 Qy 127 Pro-----HisSerProValLysGluLysProValMetSerAsnIle 140
 Db 1033 CCCTGAGGGCTTTATAAGGCTCGGGGGCCGCTGGAGCCCGCCCTCGCGCGGCCA 974
 Qy 141 GlyLysSerThrAsnIleLeuTrpHisAsnCysLeuIleGlyClnSerAspArgGlnLys 160
 Db 973 GGGCCAGAAGATCGGCCACATGGGCTCTTTCGGGGAAGACACAAAGCCCGAAGGCC 914
 Qy 161 LeuLeuGlyGlnLysGlyCysValValTrpIleThrGlyLeuSerGlySerGlyLysSer 180
 Db 913 CTTCC-GGC-----GGGGGCCAGGGGCCCTCC 888
 Qy 181 ThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTrpValLeu 200
 Db 887 TCCAGGGCCTTGGGCTTTCCAGA-----AGCCCGAACCTCTGTGGAGGAGGCTG 840
 Qy 201 AspGlyAspAsnLeuArgHis-----GlyLeuAsnArgAsp-----LeuSer 214
 Db 839 CCAAACTCAAGCCTCTCCAGAAAGCCCTAAAGCCTCTCCCGGTCCGGCTCCCGCCTTTTG 780
 Qy 215 PhelysAlaGluAspArgAlaGluAsnIleArgArgValGlyGluValAlaLysLeuPhe 234
 Db 779 CGAAGTCCACCTCCAGGGGAGGTGCGTGGCACCTTGGCCAGGTCCAGGAGAGCTTC 720
 Qy 235 AlaAspAlaGlyValIleCysIleAlaSerLeuIleSerProTrpArgArgAspArgAsp 254
 Db 719 AGATCGTCCATGTGGGCCAGGATCTCTCCCGGATGCGGGCTTCAGCCGCTCCAGGTT 660
 Qy 255 AlaCysArgAla---LeuLeuProHisSer-----AsnPheIleGluValPheIleAsp 271
 Db 659 TTGAGGAGGCTTCCAGGCTCCCGACCTCTCCAGAAAGCTTCCTCGCGCGCTCTCTCC 600
 Qy 272 LeuProLeu 274
 Db 599 ATGCCCTTG 591
 Search completed: November 2, 2002, 03:10:12
 Job time : 74 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 2, 2002, 01:53:43 ; Search time 1690 Seconds
(without alignments)
2739.323 Million cell updates/sec

Title:

US-09-720-384A-4

Perfect score: 1826

Sequence: 1 RPPHFNQTEPLVHTQPP.....PKMAKQVLCYLENGYLQA 343

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DBV=xlh
-Q=/cgn2_1/USPFO.spool/US09720384/runat_30102002_090527_19194/app_query.fasta_1.519
-DB=EST -QFMT=Fastap -SUFFIX=first -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09720384.cgn2_1.1144 -runat_30102002_090527_19194 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the total being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	936	51.3	675	10 BE357876	BE357876 DGL_22.A0
2	926	50.7	691	10 BE355111	BE355111 DGL_113.B

3	797	43.6	553	10	BI427055
4	751	41.1	536	9	AW560397
5	733.5	40.2	748	10	BG126657
6	728	39.9	564	9	AW056154
7	712	39.0	772	10	BF631366
8	708	38.8	714	10	BM407099
9	681	37.3	596	9	AW508923
10	675	37.0	576	10	BI320536
11	672.5	36.8	786	10	BI178754
12	670	36.7	602	9	AI637166
13	658	36.0	421	10	BF484142
14	604	33.1	746	10	BE361884
15	601	32.9	712	10	BG852725
16	600	32.9	681	10	BI722646
17	598	32.7	603	9	AV542370
18	597	32.7	624	10	BI995324
19	585	32.0	640	10	BI717399
20	584	32.0	721	9	AW922946
21	569	31.2	699	10	BE361874
22	555	30.4	620	9	AW330778
23	527	28.9	631	9	AW279147
24	524.5	28.7	416	9	AV424749
25	510.5	28.0	804	9	AU132232
26	508	27.8	914	9	AL552855
27	503	27.5	728	10	BE891799
28	502	27.5	722	10	BE360898
29	501.5	27.5	865	9	AL553176
30	501	27.4	739	10	BI603869
31	500	27.4	857	10	BF038478
32	497	27.2	725	10	BI693356
33	496	27.2	737	10	BI768154
34	496	27.2	803	10	BG760432
35	496	27.2	893	10	BI649921
36	495	27.1	700	10	BI685393
37	494.5	27.1	597	10	BE507893
38	494.5	27.1	727	10	BI695132
39	494.5	27.1	951	9	AL540583
40	491	26.9	760	10	BI769180
41	490.5	26.9	549	9	AVA37687
42	490.5	26.9	809	9	BE195160
43	489.5	26.8	746	9	AW957146
44	489.5	26.8	799	10	BI904228
45	489.5	26.8	860	10	BG921720

ALIGNMENTS

BE357876 BE357876 675 bp mRNA linear EST 20-JUL-2000
DGL_22.A06.gl_A002 Dark Grown 1 (DGL) Sorghum bicolor cDNA, mRNA
sequence.
BE357876 BE357876
EST. BE357876.1 GI:9299433
Sorghum.
Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Sorghum.
1 (bases 1 to 675)
Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C., Sudman,M. and Pratt
,L.H.
An EST database from Sorghum: dark-grown seedlings
Unpublished (2000)
Contact: Cordonnier-Pratt MM
Department of Botany
The University of Georgia
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 542 1805
Email: mmpratt@uga.edu
Sequences have been trimmed to exclude PolyA, vector and regions

Qy 278 GluAlaArgAspProLysGlyLeuTyrLysLeuAlaArgThrGlyLysIleLysGlyPhe 297
 Db 361 GAAGCTCGGATCCTAAGGCTGTGACAGCTTCACACAGAGAAAGATTAAGGTTTC 420
 Qy 298 ThrGlyIleAspAspProTyrGluProPheIleAsnGlyGluIleValIleLysMetLys 317
 Db 421 ACTGGAATGTGATGATCCATCAGAACCGGAGTTAATGCTGAGATGATTAATACGATGAA 480
 Qy 318 AspGluGluCysProSerProLysAlaMetAlaLysGlnValLeuCysTyrLeuGluGlu 337
 Db 481 GATGGGATGCTTACCCCAAGCAATGGCGAAGCAAGTTCTATCATACCTTGAAGAG 540
 Qy 338 AsnGlyTyrLeuGlnAla 343
 Db 541 AACGGATATTTGCAAGCT 558
 RESULT 3
 BI427055
 LOCUS
 DEFINITION sag10g11.v1 Gm-cl080 Glycine max cDNA clone GENOME SYSTEMS CLONE
 ID: Gm-cl080-718 5' similar to TR:Q9SE92 Q9SE92
 ADENOSINE-5'-PHOSPHOSULFATE KINASE ; mRNA sequence.
 BI427055
 GI:15204287
 VERSION
 KEYWORDS
 SOURCE soybean.
 ORGANISM
 Glycine max
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eustosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 Glycine.
 REFERENCE
 1. (bases 1 to 555)
 Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Coryell,V., Khanna
 ,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
 Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
 ,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
 ,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
 ,R., Waterston,R. and Wilson,R.
 Public Soybean EST Project
 Unpublished (1999)
 TITLE
 JOURNAL
 COMMENT Contact: Shoemaker R/Public Soybean EST Project
 Public Soybean EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available through: ResGen, Invitrogen Corp. 2130
 South Memorial Parkway Huntsville, AL 35801 For further information
 call: (800)-533-4363 or contact via email: ccu@resgen.com
 High quality sequence stop: 421.
 Location/Qualifiers
 1..555
 /organism="Glycine max"
 /db_xref="taxon:3847"
 /clone="GENOME SYSTEMS CLONE ID: Gm-cl080-718"
 /clone_lib="Gm-cl080"
 /tissue_type="Roots of 8 day old 'Bragg' supernodulating
 mutant N5382 seedlings"
 /dev_stage="8 days old"
 /lab_host="DH10B"
 /note="vector: pBluescript II SK+; Site_1: EcoRI; Site_2:
 XhoI; The mRNA was isolated from roots of 8 day old
 'Bragg' supernodulating mutant N5382 seedlings that were
 infected with Bradyrhizobium japonicum, strain USDA 110, 72
 hours prior to harvest. Dr. Gary Stacey generously donated
 the tissue. The roots were flash-frozen in liquid
 nitrogen. Stratagene's cDNA Synthesis Kit (catalog number
 200401) was used to synthesize the cDNA. First-strand
 synthesis was performed with 5-methyl dCTP, hence the
 ligated cDNA was hemimethylated. A modification of
 Stratagene's first-strand synthesis primer was used. An
 'anchor' nucleotide (V=A, C, or G) was added to the 3' end

of the primer (GAGAGAGAGAGAGAGAACTAGTCTCGAG(T18V) to
 anchor the primer at the 5' end of the poly(A) tract.
 After second-strand synthesis, the cDNA ends were filled
 in with cloned Pfu DNA, ligated to EcoRI adapters and
 subsequently phosphorylated. The cDNA was then
 precipitated and redissolved in sterile, RNase-,
 DNase-free water. The XhoI site within the first-strand
 synthesis primer was then restricted by digestion with
 XhoI from Promega (400/ul): all XhoI sites in the cDNA
 would be protected by their hemimethylated status. The
 cDNA constructs were size-fractionated with a 500bp
 cutoff, using Sephacryl S-500 High Resolution (Pharmacia
 Biotech) in a 2-mm diameter column and a bed volume of
 approximately 1ml. The column eluent was precipitated,
 redissolved, and ligated into Stratagene's pBluescript II
 XR predigested vector (pBluescript II SK(+)) vector that
 has been digested with EcoRI and XhoI, and phosphorylated
 by Stratagene). This library was constructed in the
 laboratory of Dr. Paul Keim and Dr. Virginia H. Coryell at
 Northern Arizona University.
 BASE COUNT 172 a 103 c 134 g 146 t
 ORIGIN

Alignment Scores:
 Pred. No.: 3.6e-44 Length: 555
 Score: 797.00 Matches: 147
 Percent Similarity: 90.27% Conservative: 20
 Best Local Similarity: 79.46% Mismatches: 18.
 Query Match: 43.65% Indels: 0
 DB: 10 Gaps: 0

US-09-720-384A-4 (1-343) x BI427055 (1-555)

Qy 157 AspArgGlnLysLeuGlyGlnLysGlyCysValValTrpIleThrGlyLeuSerGly 176
 Db 1 GAAAGGCGAAGCTACTTAACCAAAAGGAGTGTGTGTATGATTACTGACATCAGCGGA 60
 Qy 177 SerGlyLysSerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeu 196
 Db 61 TCAGGAAAAGCAGCATTTGGCATGTTCCCTAAGCAGAGACTGCACCTCAAGGGAAGTTA 120
 Qy 197 ThrTyrValLeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLys 216
 Db 121 TCTATGTCTTGTATGGAGATAACCTTCGACATGGACATAAACAAGGATCTTGTGTTTCAA 180
 Qy 217 AlaGluAspArgAlaGluAsnIleArgArgValGlyGluValAlaLysLeuPheAlaAsp 236
 Db 181 GCCGAAGATCGGCTGAAATATTTCGAGAACTGGGAAAGTGGCAAAACTGTTTGTCTGAT 240
 Qy 237 AlaGlyValIleCysIleAlaSerLeuIleSerProTyrArgArgAspArgAspAlaCys 256
 Db 241 GCTGGCTTAATATGTTGTGTAGTCTGATATCTCTTACAGAGAGACCGAGACATTCG 300
 Qy 257 ArgAlaLeuLeuProHisSerAsnPheIleGluValPheIleAspLeuProLeuLysIle 276
 Db 301 CGTGCCATCTTACCTGATGCAAAATTTATTAGGTTTTCATGAACATGCCTCTAGAAATG 360
 Qy 277 CysGluAlaArgAspProLysGlyLeuTyrLysLeuAlaArgThrGlyLysIleLysGly 296
 Db 361 TGTAGGACAGAGATCCAAAAGGCTCTATAGCTTGTCTGCGGGGGAATAATCAAGGT 420
 Qy 297 PheThrGlyIleAspAspProTyrGluProIleAsnGlyGluIleValIleLysMet 316
 Db 421 TTTACCGGATTTGATGATGCTTATGAACACCCCATAACTGTGAGATTGAAATAAAGCAG 480
 Qy 317 LysAspGluGluCysProSerProLysAlaMetAlaLysGlnValLeuCysTyrLeuGlu 336
 Db 481 GAAATGGGGATTTGCCAACCCCACTTTCAGTGGCTGGACAGTACTTACTTACTTGGAG 540
 Qy 337 GluAsnGlyTyrLeu 341
 Db 541 AATAAAGGATTTCTT 555

RESULT 4	AW560397	536 bp	nrna	linear	EST 07-SEP-2000
LOCUS	EST315445	DSIR Medicago truncatula	cdna	clone pBSIR-27A7,	nrna
DEFINITION	sequence.				
ACCESSION	AW560397				
VERSION	AW560397.1	GI:7205823			
KEYWORDS	EST.				
SOURCE	barrel medic.				
ORGANISM	Medicago truncatula				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.				
REFERENCE	1 (bases 1 to 536)				
AUTHORS	Fedorova, M., Pierson, B. L., Samac, D. A., Vance, C. P., Gantt, G. S., Peng, H., Ellis, L., Town, C. D., Bowman, C. L., Craven, M. B., Hansen, T. S., Holt, J. E. and Fraser, C. M.				
TITLE	ESTs from roots of Medicago truncatula after inoculation with <i>Phytophthora medicaginis</i>				
JOURNAL	Unpublished (1999)				
COMMENT	Contact: Carroll P. Vance Department of Agronomy and Plant Genetics University of Minnesota 411 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108 USA Tel: 612 625 5715 Fax: 651-649-5058 Email: vance004@maroon.tc.umn.edu Minnesota EST name: M251752e ; TIGR sequence name: MTBAY04TK ; More information, including clone ordering, is available at . 'http://chryslie.tamu.edu/medicago/ Seq primer: SKMOD (CTA gAA CTA gtg gAT CC).				

```

Location/Qualifiers
1. 536
/organism="Medicago truncatula"
/cultivar="genotype A17"
/db_xref="taxon:3880"
/clone="pDSIR-27A7"
/clone_lib="DSIR"
/tissue_type="roots infected with Phytophthora
medicaginis"
/dev_stage="roots harvested at 10 days post inoculation
with Phytophthora medicaginis"
/lab_host="E. coli strain XL10R"
/notes="Vector: p Bluescript SK-; Site_1: EcoRI; Site_2:
XhoI; cDNA was prepared from polyA+ enriched RNA from
roots harvested at 10 days post inoculation with
Phytophthora medicaginis. The cDNA was directionally
ligated into the Uni-ZAP XR vector from Stratagene and
packaged using Gigapack III Gold packaging extracts.
Plasmids containing cDNA inserts were excised from the
recombinant lambda-ZAP phage using Ex-Assist helper phage
and propagated in XL10R cells. Note: EST may be of fungal

```

BASE COUNT	156 a	102 c	130 g	148 t
ORIGIN	ORIGIN.			

Alignment Scores:	
Pred. No.:	4,11e-41
Score:	751.00
Length:	536
Matches:	140
Conservative:	15
Percent Similarity:	91.33%
Mismatches:	18
Best Local Similarity:	80.92%
Query Match:	41.13%
Indels:	0
DB:	9
Gaps:	0

US-09-720-384A-4 (1-343) x AW560397 (1-536)

Qy 135 ProValMetSerAsnIleGlyGlyLysSerThrAsnIleLeuTrpHisAsnCysLeuIleGly 154

Db 4 CCTCAGATGTCAAATATTGGAACTCGACGAACATTTTGTGGCATGATTGTCCAGTTCAA 63

QY 155 GlnSerAspArgGlnLysLeuLeuClyGlnLysGlyCysValValTrpIleThrGlyLeu 174
 ::: ::: ::: ::: ::: :::

Db	64	AAATGTGATAGACAGCAGCTGCTTCCACAAAGAAGTATGTATTATATGGCTAACTGGCTCTC	123
QY	175	SerGlySerGlyLysSerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGly	194
Db	124	AGTGGTTTCAGGAAAAGTACTCTTGGCATGTGCTTTGAGTCGCTTCAAGCTTGGCACTCCAGAGA	183
QY	195	HisLeuThrTyrValLeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSer	214
Db	184	AAACTGACTTACATCCTTGCAGCGTGACAATATTTCGCATGGTCTAAACCGTGATCTTACT	243
QY	215	PhelYsAlaGluAspArgAlaGluAsnIleArgArgValGlyGluValAlaLysLeuPhe	234
Db	244	TTTAGACAGACAGATCGTTTCAGAAACATACGAAAGATTGGTAGGTTGCCAAACTCTTT	303
QY	235	AlaAspAlaGlyValIleCysIleAlaSerLeuIleSerProTyrArgArgAspArgAsp	254
Db	304	GCATGCTGGTCTTATTGGCATCCTTGGATATCACCTATCAAAAGGATACAGAT	363
QY	255	AlaCysArgAlaLeuLeuProHisSerAsnPheIleGluValPheIleAspLeuProLeu	274
Db	364	GCTTGCAGAGCACTACTGCCAAGAGGGGATTTTCATTAGAGTTTCATAGACGTGCCACTA	423
QY	275	LysIleCysGluAlaArgAspProLysGlyLeuTyrLysLeuAlaArgThrGlyLysIle	294
Db	424	CATGTGTGCGAAGCTAGGAGCCCAAGGGCTATACAAGCTTGTCTGTGTCAGGAAG-ATC	482
QY	295	LysGlyPheThrGlyIleAspAspProTyrGluProPro	307
Db	483	AAAGGTTTCTACTGGTATAGATGATCCATATGAACCACT	521
RESULT 5	BG126657	748 bp	linear
LOCUS	EST472303	tomato shoot/meristem	Lycopodium obscurum cDNA clone
DEFINITION	CTOF13A21 5' sequence, mRNA sequence.		
ACCESSION	BG126657		
VERSION	BG126657.1	GI:12626845	
KEYWORDS	EST.		
SOURCE	tomato.		
ORGANISM	Lycopodium obscurum		
REFERENCE	Eukaryota; Viridiplantae;		
AUTHORS	van der Hoeven R., Bezzlerides J., Sun, H., Cho, J., Utterback, T., Hansen C., Renning C. and Tanksley, S.		
TITLE	Generation of ESTs from tomato shoot/meristem tissue		
JOURNAL	Unpublished (2001)		
COMMENT	Contact: CUGI		
	1. (bases 1 to 748)		
	van der Hoeven R., Bezzlerides J., Sun, H., Cho, J., Utterback, T., Hansen C., Renning C. and Tanksley, S.		
	Generation of ESTs from tomato shoot/meristem tissue		
	Unpublished (2001)		
	Contact: CUGI		
	Clemson University Genomics Institute		
	Clemson University		
	100 Jordan Hall, Clemson, SC 29634, USA		
	Email: http://www.genome.clemson.edu/orders/index.html .		

[illegible]

1. .748

/organism="Lycopersicon esculentum"

/cultivar="TA496"

```
/db_xref="taxon:40
```

```
/clone="CTOF13A21"
```

```
/clone_lib="tomato shoot/meristem"
```

```

/tissue_type="shoot/meristem"
/sex_stage="developing shoots
from 1-6 y old plants"

```

```
/dev_stage="developing shoots from 4-6wks old plants"
/lab_host="SOLR"
```

```

/lab_host="SOLR"
/note="vector: pB|uescript SK(-): site 1: EcoR1: site 2:

```

```

/note=vector; ptnescript sk(-); site_1: ecor1; site_2:
Xhol: small expanding leaves from the growing tip were

```

small expanding leaves from the growing tip were taken from greenhouse plants (4-6wks old TA496). Tissue

was immediately frozen in liquid nitrogen and stored at -80°C until analysis.

227 a

ORIGIN

Alignment Scores:

Alignment scores:	
Pred. No.:	9.25e-40
Length:	748

748

Score: 733.50 Matches: 142
 Percent Similarity: 79.62% Conservative: 26
 Best Local Similarity: 67.30% Mismatches: 36
 Query Match: 40.17% Indels: 7
 DB: 10 Gaps: 1

US-09-720-384A-4 (1-343) x BG126657 (1-748)

Qy 136 ValMetSerAsnIleGlyLysSerThrAsnIleLeuTrpHisAsnCysLeuIleGln 155
 Db 119 TTGATCTCTACTGTGGCAATCGGCAACATATCTTGGCATGAAATCCAGTTGGGAAG 178
 Qy 156 SerAspArgGlnLysLeuGlyGlnLysGlyCysValValTrpIleThrGlyLeuSer 175
 Db 179 CGCGAAGGAAAGCTGCTTACCAACAGGATGTGTGTATGATCAGCGGTCTCAGT 238
 Qy 176 GlySerGlyLysSerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHis 195
 Db 239 GGATCAGGAAAGCACATCTGATCTTCCCTAGTAGAGAGTTGCGAGTCAAGGCTAAG 298
 Qy 196 LeuThrTyrValLeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPhe 215
 Db 299 CTTTCATAGCTTGTGTGATGTCACACCTTCGSCATGCTCTGAAACAGAAATCTTGGGTC 358
 Qy 216 LysAlaGluAspArgAlaGluAsnIleArgArgValGlyGluValAlaLysLeuPheAla 235
 Db 359 TCACCAAGAAAGCGCACTGAGATATACGTAGGACTGGGGAAGTTGCAAAATCTCTTGCA 418
 Qy 236 AspAlaGlyValIleCysIleAlaSerLeuIleSerProTyrArgArgAspAla 255
 Db 419 GATGCTGGATTAATTTGCTTGCATGCAAGTTTGATATCTCTTACAGAAAGGATGCTGATGCT 478
 Qy 256 CysArgAlaLeuLeuProHisSerAsnPheIleGluValPheIleAspLeuProLeuLys 275
 Db 479 TGGCGTGCAATATCCCAAGATAAAGTTCAATGAGGTTTTTATGAATATGCTCTACAA 538
 Qy 276 IleCysGluAlaArgAspProLysGlyLeuTyrLysLeuAlaArgThrGlyLysIleLys 295
 Db 539 CTGTGTGAAGATACAACTAAAGACGGAGTAGTTCCTACACCATCATGAGTGGGGGCA 598
 Qy 296 -----GlyPheThrGlyIleAspAspProTyrGluProIleAsnGly 310
 Db 599 GCACCTCATTTGTAGTTTACTGGATAGATGATCTTATGACCACTTTGAATGCT 658
 Qy 311 GluIleValIleLysMetLysAspGluGlyCysProSerProLysAlaMetAla-LysGI 330
 Db 659 GAGATTGAATACAACTAAAGACGGAGTAGTTCCTACACCATCATGAGTGGGGGCA 718
 Qy 330 nValLeu-CystyrLeuGluGluAsnGly 339
 Db 719 AGTAGTTTCTTATATGAGAAATGAAGG 747

RESULT 6
 AW056154
 LOCUS 600004H03.y1 660 - Mixed stages of anther and pollen Zea mays cDNA,
 DEFINITION mRNA sequence.
 ACCESSION AW056154
 VERSION AW056154.1 GI:5928862
 KEYWORDS EST.
 SOURCE Zea mays.

ORGANISM Zea mays.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 564)
 Walbot.V.

REFERENCE
 AUTHORS Maize ESTs from various cDNA libraries sequenced at Stanford
 TITLE University
 JOURNAL Unpublished (1999)
 COMMENT Contact: Walbot V
 Department of Biological Sciences
 Stanford University

855 California Ave., Palo Alto, CA 94304, USA
 Tel: 650 723 2227
 Fax: 650 725 8221
 Email: walbot@stanford.edu
 Plate: 660004 row: H column: 03.

FEATURES
 source

1..564
 /organism="Zea mays"
 /cultivar="Ohio43"
 /db_xref="taxon:4577"
 /clone_lib="660 - Mixed stages of anther and pollen"
 /tissue_type="whole premeiotic anthers to pollen shed"
 /dev_stage="premeiotic anthers to pollen shed"
 /lab_host="XLOUP"
 /note="Organ: anthers; Vector: Lambda Zap; Site_1: EcoRI;
 Site_2: XhoI; Anther and pollen cDNA library.
 Directionally sequenced with 5' end at the EcoRI site.
 Created by Amie Franklin."
 BASE COUNT 124 a 166 c 170 g 102 t
 ORIGIN

Alignment Scores:
 Pred. No.: 1,51e-39 Length: 564
 Score: 728.00 Matches: 139
 Percent Similarity: 88.95% Conservative: 14
 Best Local Similarity: 80.81% Mismatches: 19
 Query Match: 39.87% Indels: 0
 DB: Gaps: 0

US-09-720-384A-4 (1-343) x AW056154 (1-564)

Qy 171 IleThrGlyLeuSerGlySerGlyLysSerThrLeuAlaCysAlaLeuSerArgGluLeu 190
 Db 7 ATCACTGGCTTAAAGCGTTTCAGGAAAAAGCAACGCTCGCTGGCGCTGAGCGCGAGCTG 66
 Qy 191 HisCysArgGlyHisLeuThrTyrValLeuAspGlyAspAsnLeuArgHisGlyLeuAsn 210
 Db 67 CACGGCAGAGGCGACCTCAGCTAGTCTCGAGCGGCAACCTCAGCAGCGCGCTGAAC 126
 Qy 211 ArgAspLeuSerPheLysAlaGluAspArgAlaGluAsnIleArgArgValGlyGluVal 230
 Db 127 AGGACCTCAGCTTCGGAGCAGAGACCGCGCGAGAACATNCCGAGTAGGGGAAGNA 186
 Qy 231 AlalysLeuPheAlaAspAlaGlyValIleCysIleAlaSerLeuIleSerProTyrArg 250
 Db 187 GCGAAGCTGTTCGCGAGCTGGCTCTGCTGTCGCGAGCGCCCAAGGCTCTACAGCTCGCACGC 246
 Qy 251 ArgAspArgAspAlaCysArgAlaLeuLeuProHisSerAsnPheIleGluValPheIle 270
 Db 247 AGCGACCGAAGCGCGTGTGCGGATCTGCTGCCCAAGCACTCGTTTATCGAGGTGTTCTTG 306
 Qy 271 AspleuProLeuLysIleCysGluAlaArgAspProLysGlyLeuTyrLysLeuAlaArg 290
 Db 307 GAGTCGCGCTCAAGTGTGGAAGCGGAGCGGACCCCAAGGCTCTACAGCTCGCACGC 366
 Qy 291 ThrGlyLysIleLysGlyPheThrGlyIleAspAspProTyrGluProIleAsnGly 310
 Db 367 GCGCGCAAAATCAAAAGGTTTCCCGGCATCGAGCATCTTACGAACCGCGCTCGGACTGT 426
 Qy 311 GluIleValIleLysMetLysAspGluGluCysProSerProLysAlaMetAlaLysGln 330
 Db 427 GAGATAGTATCATCGATGTAAGTCGGCGCTGCCCTTCGCTGAATCATGATGCTGTGTCAC 486
 Qy 331 ValLeuCysTyrLeuGluAsnGlyTyrLeuGln 342
 Db 487 GTTGTGTCGTACTTTCAGACGAATGGTTTCTCCACG 522

RESULT 7
 BF631366 772 bp mRNA linear EST 22-OCT-2001
 LOCUS HVSMB0015L17f Hordeum vulgare seedling shoot EST library
 DEFINITION HVCNDA0002 (Dehydration stress) Hordeum vulgare cDNA clone
 HVSMB0015L17f, mRNA sequence.

ACCESSION BF631366
 VERSION BF631366.2 GI:13092061
 KEYWORDS EST.
 SOURCE Hordeum vulgare
 ORGANISM Hordeum vulgare
 REFERENCE 1 (bases 1 to 772)
 AUTHORS Wing, R., Close, T.J., Kleinohfs, A., Wise, R., Begum, D., Frisch, D., Yu, Y., Henry, D., Palmer, M., Rambo, T., Simmons, J., Choi, D.W., Fenton, R.D., Oates, R. and Main, D.
 TITLE Development of a genetically and physically anchored EST resource for barley genomics: Morex drought-stressed seedling shoot cDNA library
 JOURNAL Unpublished (2001)
 COMMENT On Dec 19, 2000 this sequence version replaced gi:11895524.
 Contact: Wing RA
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Tel: 864 656 7288
 Fax: 864 656 4293
 Email: rwing@clemson.edu
 Total hq bases = 462
 Seq primer: ATTACCTCCTCAATAAGGG
 High quality sequence stop: 675.
 Location/Qualifiers
 1. .772
 /organism="Hordeum vulgare"
 /cultivar="Morex"
 /db_xref="taxon:4513"
 /clone_lib="HVSMEB0015L17f"
 /HVCdna0002 (Dehydration stress)
 /tissue_type="Seedling shoot"
 /lab_host="fJCL21"
 /note="Vector: lambda2AP; Site_1: EcoRI; Site_2: XhoI;
 Seeds were surface sterilized then germinated under axenic conditions in the dark at room temperature on filter paper with water, nystatin and ceftaxime in covered crystallization dishes. Five-day old seedlings were incubated at 90% RH for 24 hr. Shoots were then harvested, total RNA was prepared, poly(A) RNA was purified, one primary unamplified cDNA library was made, 600000 pfu were in vivo excised to give pBluescript SK(-) cDNA phagemids. These steps were performed in the TJ Close laboratory at the University of California, Riverside (Choi, Close, Fenton). Phagemids were plated and picked at the Clemson University Genomics Institute (CUGI) (Begum, Palmer, Frisch, Atkins and Wing). Plasmid DNA preparations, DNA sequencing and sequence analysis were performed at CUGI (Wing, Yu, Frisch, Henry, Simmons, Oates, Rambo, Main). The sequence has been trimmed to remove vector sequence and contains a minimum of 100 bases of phred value 20 or above. For more details on library preparation and sequence analysis see <http://www.genome.clemson.edu/projects/barley>. To order this clone see <http://www.genome.clemson.edu/orders> Also see Close TJ, Wing R, Kleinohfs A, Wise R (2001) Genetically and physically anchored EST resources for barley genomics. Barley Genetics Newsletter 31:29-30. (<http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html>)"

BASE COUNT 202 a 146 c 238 g 186 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 2,64e-38 Length: 772
 Score: 712.00 Matches: 140
 Percent Similarity: 79.31% Conservative: 21
 Best Local Similarity: 68.97% Mismatches: 41
 Query Match: 38.99% Indels: 2
 DB: 10 Gaps: 0

US-09-720-384A-4 (1-343) x BF631366 (1-772)

Qy 138 SerAsnIleGlyLysSerThrAsnIleLeuTrpHisAsnCysLeuIleGlyGlnSerAsp 157
 Db 72 TCAACTGTGCCAAGATCATCAATATATCTTCTGCATGATGTCGCCAGTGGGCAAGACTGAC 131
 Qy 158 ArgGlnLysLeuGlyGlnLysGlyCysValValTTPileThrGlyLeuSerGlySer 177
 Db 132 CGCAGAAATCTACTACGACAGAAAGGGTGTGTGTGTGGATTACAGGCTTAGTGTTC 191
 Qy 178 GlyLysSerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThr 197
 Db 192 GGTAAAGTACCTGGCATGTCACATTAGTCTCAGAGCTCCATACAGAGGGAAGCTGGC 251
 Qy 198 TyrValLeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLysAla 217
 Db 252 TATGTTCTTGATGGCGATACTTAAGACATGGTCTTAACAAGGATCTTGGGTTCCGAGCT 311
 Qy 218 GluAspArgAlaGluAsnIleArgArgValGlyGluValAlaLeuLysLeuPheAlaAspAla 237
 Db 312 GAAGATCGTGTGCTGAAATATACGAGAGTTGGTGAAGTTGCCAAGCTATTTCGAGATGCA 371
 Qy 238 GlyValIleCysIleAlaSerLeuIleSerProTyrArgArgAspArgAspAlaCysArg 257
 Db 372 GGTCTAGTGTGCATTGCTAGTTTATATCTCCATATAGGAGAGACCGAGAGCTCTGTCTGT 431
 Qy 258 AlaLeuLeuProHisSerAsnPheIleGluValPheIleAspLeuProLeuLysIleCys 277
 Db 432 GCACCTGTTCTCAGAGGTAGTTTATTGAAGTTGCTTGAACATGTCCTTGGAAATGTGT 491
 Qy 278 GluAlaArgAspProLysGlyLeuTyrLysLeuAlaArgThrGlyLysIleLysGlyPhe 297
 Db 492 GAAGCAAGGATGCCAAGGGCTTTACAGCTTGTCTGCAAGAAAAAAT-AAAGGGTTT 550
 Qy 298 ThrGlyLeuAspAspProTyrGluProProIleAsnGlyGluIle-ValIleLysMetly 317
 Db 551 ACTGGTGTGTGATGCCCATATGAAGCCCATGGAATTCGAGATTCGCCATCAGGAGCT 610
 Qy 317 sAspGluGlyCysProSerProLysAlaMetalalysGlnValLeuLeuCysTyrLeuGlu 337
 Db 611 GGATGCGTGTGCCCTTAGCCTTCGACTTGGGAACACAAAGTATGATGCTGCTGCGGA 670
 Qy 337 uAsnGly 339
 Db 671 GGGGGGG 677
 RESULT 8
 BM407099
 LOCUS EST581426 potato roots Solanum tuberosum cDNA clone CP029H6 5' end
 DEFINITION , mRNA sequence.
 ACCESSION BM407099
 VERSION BM407099.1 GI:18258729
 KEYWORDS EST.
 SOURCE potato.
 ORGANISM Solanum tuberosum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
 REFERENCE 1 (bases 1 to 714)
 AUTHORS van der Hoeven, R., Sun, H., Karamycheva, S.A., Tsai, J., Van Aken, S., Uterback, T., Chiemingo, A., Bougri, O., Bougri, C.R., Ronning, C., Tanksley, S. and Baker, B.
 TITLE Generation of ESTs from potato roots
 JOURNAL Unpublished (2001)
 COMMENT Contact: Research Genetics, Libraries Division
 Tel: 1-800-711-6195
 Email: cdna@resgen.com
 For clone info: please contact Research Genetics, Libraries Division tel 1-800-711-6195, email cdna@resgen.com
 Seq primer: T3
 FEATURES Location/Qualifiers


```
source
1. .714
/organism="Solanum tuberosum"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="cPR029H6"
/tissue_type="roots"
/dev_stage="in vitro grown stem cuttings"
/lab_host="SOLR"
/notes="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; supplier: Cornell University, Tanksley lab;
sequencing: The Institute for Genomic Research. Roots were
isolated from in vitro grown stem cuttings on CM medium.
Roots were isolated two weeks after placing the stem
cuttings from in vitro grown plants on medium."

BASE COUNT 197 a 126 c 172 g 219 t
ORIGIN

Alignment Scores:
Pred. No.: 4,43e-38 Length: 714
Score: 708.00 Matches: 129
Percent Similarity: 85.23% Conservative: 21
Best Local Similarity: 73.30% Mismatches: 26
Query Match: 38.77% Indels: 0
DB: 10 Gaps: 0

US-09-720-384A-4 (1-343) x BM407099 (1-714)
Qy 136 ValMetSerAsnIleGlyLysSerThrAsnIleLeuTrpHisAsnCysLeuIleGlyGln 155
Db 56 TTGATGCTCTACTGTGGCAATCGGCAACATATCTTGGCATGATAATCAGTTGGGAAG 115
Qy 156 SerAspArgGlnLysLeuGlyGlnLysGlyCysValValTrpIleThrGlyLeuSer 175
Db 116 GTGGAAGGAAAGAGCTGCTTACCACAGGGGTGTGTGTATGATGATCAGGTCTCAGT 175
Qy 176 GlySerGlyLysSerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHis 195
Db 176 GGATCAGGAAAGGAGCGCTTGATCTCCCTAGGTAGAGAAATGTCAGTCAAGGGGTAG 235
Qy 196 LeuThrTyrValLeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPhe 215
Db 236 CTTTCATAGCTTCTGTGTGACACCTTAGCATGCTCGACACAGAAATCTTGGGTTC 295
Qy 216 LysAlaGluAspArgAlaGluAsnIleArgArgValGlyGluValAlaLysLeuPheAla 235
Db 296 TCACCAGAAAGCGGACTGAGAATATACGTAGGACTGGGAAGTTCGAATCTCTTTGCA 355
Qy 236 AspAlaGlyValIleCysIleAlaSerLeuIleSerProTyrArgArgAspArgAspAla 255
Db 356 GATGCTGGATTAATTTGCATTCAGTTTGATATCTCTTACAGAAAGGATCGTATGCT 415
Qy 256 CysArgAlaLeuLeuProHisSerAsnPheIleGluValPheIleAspLeuProLeuLys 275
Db 416 TGCCGCGCATTAATTTGCCAGATAAAAGTTTCATTGAGGTTTATGAAATATGCTCTACAA 475
Qy 276 IleCysGluAlaArgAspProLysGlyLeuTyrLysLeuAlaArgThrGlyLysIleLys 295
Db 476 CTGTGTGAAGATAGAGATCCAAAAGGCGCTCTACAAGCTAGCTCTGAGGAGTAAATCAA 535
Qy 296 GlyPheThrGlyIleAspAspProTyrGluProProIleAsnGlyGlu 311
Db 536 GGTGTTTACGGAGTAGATGATCTCTTATGAACCACTTTGTAATTTGTGAG 583

RESULT 9
AW508923
LOCUS
DEFINITION
s138a06.v1 Gm-r1030 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
Gm-r1030-1163 5' similar to TR:049204 049204
ADENOSINE-5'-PHOSPHOSULFATE-KINASE ;, mRNA sequence.
ACCESSION
AW508923
VERSION
AW508923.1 GI:7147001
KEYWORDS
EST.
```

```
soybean.
Glycine max
Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 596)
Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Coryell,V., Khanna
,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
,R., Waterson,R. and Wilson,R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130
South Memorial Parkway Huntisville, AL 35801 For further information
Call: (800) 533-4363 or contact via email: ccu@resgen.com
Insert Length: 1061 Std Error: 0.00
High quality sequence stop: 444.
Location/Qualifiers
1. .596
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-r1030-1163"
/clone_lib="Gm-r1030"
/lab_host="DH10B"
/notes="Vector: pSPOR1; Site_1: SalI; Site_2: NotI; This
cDNA library was constructed from mRNA isolated from
immature cotyledons of greenhouse grown plants
(individual seed fresh weight of 100-300mg). The library
was prepared using the Life Technologies pSuperScript cDNA
library construction kit. Complementary DNA was
synthesized from mRNA using a poly(dT) sequence with a
NotI restriction site. SalI linkers adapters were ligated
to the blunt-ended cDNA fragments followed by NotI
digestion. The cDNA fragments were directionally cloned
into the NotI-SalI restriction site of the pSPOR1
vector. The ligated cDNA fragments were transformed into
E. coli ElectroMax DH10B host cells. This library was
constructed by Dr. Lila Vodkin and Dr. Anu Khanna. Note
that Gm-r1030 is a re-rack of Gm-cl007."
BASE COUNT 182 a 119 c 131 g 163 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 2,25e-36 Length: 596
Score: 681.00 Matches: 125
Percent Similarity: 89.31% Conservative: 17
Best Local Similarity: 78.62% Mismatches: 17
Query Match: 37.29% Indels: 0
DB: 9 Gaps: 0

US-09-720-384A-4 (1-343) x AW508923 (1-596)
Qy 137 MetSerAsnIleGlyLysSerThrAsnIleLeuTrpHisAsnCysLeuIleGlyGlnSer 156
Db 120 ATGGCAACTCTGAGCAATTCACATTAATATCTTTTGGCAAGATGCTCAATAGGAGCGCT 179
Qy 157 AspArgGlnLysLeuGlyGlnLysGlyCysValValTrpIleThrGlyLeuSerGly 176
Db 180 GAAAGCGCAAGACTACTTAACCAAGGGATGTGTGTATGATGATGACTGACGCGCA 239
Qy 177 SerGlyLysSerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeu 196
Db 240 TCAGGAAAAAGCACATTGGCATGCTCCCTAAGCAGAGAACTGCATCAAGGGGAAAGTTA 299
```

197 ThrTyrValLeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLys 216
 Db 300 TCTTATGTCCTTGTATGAGATAACCTTCGGCATGACTAAACAAGATCTTGGTTTAA 359
 QY 217 AlaGluAspArgAlaGluAsnIleArgValGlyGluValAlaLysLeuPheAlaAsp 236
 Db 360 GCTGAAGATCGCACTGAAATATTCGAGAACTGCAGAACTCTTCGTCAT 419
 QY 237 AlaGlyValIleCysLeuAlaSerLeuLeuSerProTyrArgArgAspArgAspLys 256
 Db 420 GCTGGTTTAATATGTTCTAGTCTGATATCTCTTACAGAAAGATCGAGACATGC 479
 QY 257 ArgAlaLeuLeuProHisSerAsnPheIleGluValPheIleAspLeuProLeuLysIle 276
 Db 480 CGTGCCATGTTACCTGATGCAAAATTTATTGAGGTTTTTCATGAACATGCTCTAGAATG 539
 QY 277 CysGluAlaArgAspProLysGlyLeuTyrLysLeuAlaArgThrGlyLysIleLys 295
 Db 540 TCGGAGCAGAGATCAAAAGGCCTCTATAAGCTTCGTCGANGAAATCAAA 596

RESULT 10
 BI320536
 LOCUS
 DEFINITION
 sahs6h02.y1 Gm-c1049 Glycine max cDNA clone GENOME SYSTEMS CLONE
 ID: Gm-c1049-2643 5', similar to SW:KAPS.CATRO 049204
 ADENYLISULFATE KINASE, CHLOROPLAST PRECURSOR ;, mRNA sequence.

ACCESSION
 BI320536
 VERSION
 BI320536.1 GI:14999722
 KEYWORDS
 EST.
 SOURCE
 soybean.
 ORGANISM
 Glycine max

REFERENCE
 AUTHORS
 Shoemaker,R., Kelm,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna
 A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
 Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
 Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
 R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
 R., Waterston,R. and Willson,R.
 Public Soybean EST Project
 Unpublished (1999)
 TITLE
 JOURNAL
 COMMENT
 Contact: Shoemaker R/Public Soybean EST Project
 Public Soybean EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available through: ResGen, Invitrogen Corp. 2130
 South Memorial Parkway Huntsville, AL 35801 For further information
 call: (800)-533-4363 or contact via email: ccu@resgen.com
 Seq primer: -40RP from Gibco
 High quality sequence stop: 440.
 Location/Qualifiers
 1. .576
 /organism="Glycine max"
 /db_xref="taxon:3847"
 /clone="GENOME SYSTEMS CLONE ID: Gm-c1049-2643"
 /clone_lib="Gm-c1049"
 /tissue_type="whole seedlings of greenhouse grown plants"
 /dev_stage="3 week old"
 /lab_host="DH10B"
 /note="Vector: pBluescript II SK+; Site:1: EcoRI; Site:2:
 XhoI; The Clark NIL was constructed and seed was provided
 by Dr. J. Specht, University of Nebraska (Shoemaker and
 Specht, 1995). The cDNA library was constructed from mRNA
 isolated from whole seedlings of 3 week old greenhouse
 grown plants. Complementary DNA was synthesized from mRNA
 using a primer consisting of a poly(dT) sequence with a

XhoI restriction site and a 3' anchor. EcoRI adapters were
 ligated to the blunt-ended cDNA fragments followed by XhoI
 digestion. The cDNA fragments were directionally cloned
 into the EcoRI-XhoI restriction site of the pBluescript
 vector. The ligated cDNA fragments were transformed into
 DH10B host cells (GibcoBRL). The library was constructed
 in cooperation with Dr. Paul Kelm's laboratory at Northern
 Arizona University."

BASE COUNT 185 a 100 c 130 g 161 t
 ORIGIN

Alignment Scores:
 Pred. No.: 5,44e-36 Length: 576
 Score: 675.00 Matches: 122
 Percent Similarity: 89.09% Conservative: 25
 Best Local Similarity: 73.94% Mismatches: 18
 Query Match: 36.97% Indels: 0
 DB: 10 Gaps: 0

US-09-720-384A-4 (1-343) x BI320536 (1-576)

QY 131 ValLysGluLysProValMetSerAsnIleGlyLysSerThrAsnIleLeuTrpHisAsn 150
 Db 43 ATAAAGGTGACTTTTCTGATGTCAAATGTGGCAACTCGACAAATATTATGTGGCATGAC 102
 QY 151 CysLeuIleGlyGlnSerAspArgGlnLysLeuLeuGlyGlnLysGlyCysValValTrp 170
 Db 103 TGCTAAATTAGAAATAGATACAGACAGCTGCTTCAGCAAAAGGCTGCTTTATATGG 162
 QY 171 IleThrGlyLeuSerGlySerGlyLysSerThrLeuAlaCysAlaLeuSerArgGluLeu 190
 Db 163 CTAATGCGCCACAGTGGTTCAGAAACCACTATTGTCATGCTGTGATCGCAAGCTTG 222
 QY 191 HisCysArgGlyHisLeuThrTyrValLeuAspGlyAspAsnLeuArgHisGlyLeuAsn 210
 Db 223 CACTCCAAAGAAACCTGCTTACATCCTTGATGTCACATATTTCGGCATGCTCTAAAC 282
 QY 211 ArgAspLeuSerPheLysAlaGluAspArgAlaGluAsnIleArgValGlyGluVal 230
 Db 283 CAAGATCTTAGTTTCAGAGAGAGATCGTCTGAAACATATTAGAGATTTGGTGGAGTG 342
 QY 231 AlaLysLeuPheAlaAspAlaGlyValIleCysIleAlaSerLeuIleSerProTyrArg 250
 Db 343 GCTAAACTCTTTCAGATGCTGCTGTTATTTCATCCTGATCCTAGTTTAATATCACCATACCAA 402
 QY 251 ArgAspArgAspAlaCysArgAlaLeuLeuProHisSerAsnPheIleGluValPheIle 270
 Db 403 AAGGATAGAGATCCATCGCAGACACTAATGCCAAAGGAGATTTATTGAGTTTTCATA 462
 QY 271 AspLeuProLeuLysIleCysGluAlaArgAspProLysGlyLeuTyrLysLeuAlaArg 290
 Db 463 GATGTTCCACTCATGTTGTGTAAGCTACGAGCCCAAGGAGGCTCTACAAGCTTGCTCGA 522
 QY 291 ThrGlyLysIleLys 295
 Db 523 GCTGGAAAGATCAAA 537

RESULT 11
 BI178754
 LOCUS
 DEFINITION
 EST519699 cSTE Solanum tuberosum cDNA clone cSTE15J16 5' sequence,
 mRNA sequence.

ACCESSION
 BI178754
 VERSION
 BI178754.1 GI:14644565
 KEYWORDS
 EST.
 SOURCE
 potato.
 ORGANISM
 Solanum tuberosum

REFERENCE
 AUTHORS
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
 1 (bases 1 to 786)
 van der Hoeven,R., Bezzerides,J., Bachem,C., Visser,R., Cho,J.,
 Chiemingo,A., Bougri,O., Buell,C.R., Romming,C., Tanksley,S. and

Baker.B.
Generation of ESTs from in vitro grown microtubers
Unpublished (2001)
Contact: Cathy Roming
The Institute for Genomic Research
For clone info: please contact Research Genetics, Libraries
Division tel 1-800-711-6195, email cdm@resgen.com
Seq primer: M13F-R.

FEATURES
source
Location/Qualifiers
1..786
 /organism="Solanum tuberosum"
 /cultivar="Bintje"
 /db_xref="taxon:4113"
 /clone_lib="CSTEL516"
 /clone_lib="CSTE"
 /tissue_type="axillary buds of stem explants; growing
sink-tubers"
 /dev_stage="7, 8 and 10 days"
 /lab_host="SOLR"
 /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; Tissue supplied by Christian Bachem and Richard
Visser (Department of Plant Breeding, Wageningen
University, The Netherlands). The cDNA libraries will
attempt to capture the induction and initiation/initial
growth of the tuber in an in vitro system as described in
Bachem et al. (Plant Journal, 1996). Small microtubers
develop from axillary buds attached to stem explants when
placed on a high sucrose medium (10%). Visible
morphological changes occur synchronously at day five in
the axillary buds. The first library, cSTA (1-20) consists
of axillary buds harvested on days 1-3. This targets
those genes involved in induction of the microtubers.
Following libraries, cSTA (21-40) and cSTA (41-60),
capture genes involved in tuber initiation and outgrowth.
This library is noted as P3 in Rankley lab notebooks."

BASE COUNT 212 a 137 c 219 g 218 t

ORIGIN

Alignment Scores:

Pred. No.:	1.18e-35	Length:	786
Score:	672.50	Matches:	130
Percent Similarity:	87.65%	Conservative:	19
Best Local Similarity:	76.47%	Mismatches:	21
Query Match:	36.83%	Indels:	2
DB:	10	Gaps:	0

US-09-720-384A-4 (1-343) x BI178754 (1-786)

Qy	135	ProValMetSerAsnIleGlyLysSerThrAsnIleLeuTrpHisAsnCysLeuIleGly	154
Db	278	CCTCAGATGCTTCAAAATAGGGAAATTCACAAACATTGTGGCACAGTGCTCTAGAG	337
Qy	155	GlnSerAspArgGlnLysLeuLeuGlyLntLysGlyCysValValTrpIleThrGlyLeu	174
Db	338	AAATGTGCACAGAGAGGATTGCTTCAACAACGGGGTTGTATATGGGTGAACGGGTCTC	397
Qy	175	SerGlySerGlyLysSerThrLeuAlaCysAlaLeuSerArgGlnLeuHisCysArgGly	194
Db	398	AGTGGTTCAGGGAAGAGACTCTCGCATGTGCTTAGTCTGGCTTACACCCTAGAGA	457
Qy	195	HisLeuThrTrpValLeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSer	214
Db	458	AAACTCACCTATATCCCTTGATGGTGATAATGTGCCACTGTCTGAATCGTATCTTAGT	517
Qy	215	PheLysAlaGluAspArgAlaGluAsnIleArgArgValGlyGluValAlalysLeuPhe	234
Db	518	TTCAGACGACAGACCCGACGACAGATATACACGGACTGGGAGGTGGCAAGCTTTTC	577
Qy	235	AlaAspAlaGlyValIleCysIleAlaSerLeuIleSerProTyrrArgArgAspArgasp	254
Db	578	GCGATGCTGGAGTATTTTCATTTGCCAGTTTTGATCTCCCTATAGAAAGAACGACGAT	637
Qy	255	AlaCysArgAlaLeuLeuProHisSerAsnPhelIleGluValPheIleAspLeuProLeu	274

```

QY 243 AlaSerLeuIleSerProTyrArgArgAspArgAspAlaCysArgAlaLeuLeuProHis 262
|||||
Db 421 GCACGCCATATCGCCCTACAGACGACCGCGGTGTCGCATCTGTCGCCAAG 362
|||||
QY 263 SerAsnPhelIleGluValPheIleAspLeuProLeuLysIleCysGluAlaArgAspPro 282
|||||
Db 361 CACTCGTTATCGAGGTGTTCTCTGGACGTGCGCTTCAAGTGTGCCAAGCCAGGACCCC 302
|||||
QY 283 LysGlyLeuTyrLysLeuAlaArgThrGlyLysIleLysGlyPheThrGlyIleAspAsp 302
|||||
Db 301 AAAGGCCCTCTACAAGTCGACGCGCGCAAAATCAAGGGTTCACCGGCATCCAGCAT 242
|||||
QY 303 ProTyrGluProProIleAsnGlyGluIleValIleLysMetLysAspGluGluCysPro 322
|||||
Db 241 CCTTACGAACCGCGCTGGACTGTGAGATAGTATCCAGTGTAAAGTCGCGACATGCCCT 182
|||||
QY 323 SerProLysAlaMetAlaLysGlnValLeuCysTyrLeuGluGluAsnGlyTyrLeuGln 342
|||||
Db 181 TCGCCTGAATCGATRAAAAAAACGTTGTGCTGCTACCTTGAGACGAATGGTTTCCTCCAG 122
|||||
RESULT 13
BF484142 421 bp mRNA linear EST 06-DEC-2000
LOCUS WHEI1788_G10_M20ZS wheat pre-anthesis spike cDNA library Triticum
DEFINITION aestivum cDNA clone WHEI1788_G10_M20, mRNA sequence.
ACCESSION BF484142
VERSION BF484142.1 GI:11567443
KEYWORDS EST.
SOURCE bread wheat.
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Triticum.
REFERENCE 1 (bases 1 to 421)
AUTHORS Anderson,O.D., Chao,S., Choi,D.W., Close,T.J., Fenton,R.D., Han
,P.S., Hsia,C.C., Kang,Y., Lazo,G.R., Miller,R., Rausch,C.J.,
Seaton,C.L. and Tong,J.C.
The structure and function of the expressed portion of the wheat
genomes - Pre-anthesis spike cDNA library
Unpublished (2000)
Contact: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific
West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105959773
Fax: 5105959818
Email: oanderson@wv.usda.gov
Sequence have been trimmed to remove vector sequence and low
quality sequence with phred score less than 20
Seq primer: Stratagene SK primer.
FEATURES
Location/Qualifiers
source 1..421
/organism="Triticum aestivum"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="WHEI1788_G10_M20"
/clone_lib="wheat pre-anthesis spike cDNA library"
/tissue_type="Spike before anthesis"
/dev_stage="Adult plant"
/lab_host="E. coli SOLR"
/notes="Vector: Lambda Uni-ZAP XR, excised phagemid;
Site_1: EcoRI; Site_2: XhoI; Plants were grown in the
greenhouse. Whole spike with awns trimmed, white, green
and yellow anther were collected and total RNA, and
poly(A) RNA were prepared, a cDNA library was made, and
the cDNA clones were in vivo excised to give pBluescript
phagemids in the TJ Close lab (Choi, Close, Fenton) at
the University of California, Riverside. plasmid DNA
preparations and DNA sequencing were performed in the OD
Anderson lab (all other authors)."
BASE COUNT 126 a 84 c 102 g 108 t 1 others
ORIGIN

```

```

Alignment Scores:
Pred. No.: 5,02e-35 Length: 421
Score: 658.00 Matches: 124
Percent Similarity: 93.57% Conservative: 7
Best Local Similarity: 88.57% Mismatches: 9
Query Match: 36.04% Indels: 0
DB: 10 Gaps: 0

US-09-720-384A-4 (1-343) x BF484142 (1-421)
QY 142 LysSerThrAsnIleLeuTyrPheHisAsnCysLeuIleGlyGlnSerAspArgGlnLysLeu 161
|||||
Db 1 AAATCAACCTAATATTCTATGCGCATGACTGCCCAATTGGGCAACCTCGACACAGAAATTTG 60
|||||
QY 162 LeuGlyGlnLysGlyCysValIleThrGlyLeuSerGlySerGlyLysSerThr 181
|||||
Db 61 CTGGGGCAAAAAGGATGTCTATATGATCACAGGACTCAGCGGTTTCAGGGAAGTACC 120
|||||
QY 182 LeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTyrValLeuAsp 201
|||||
Db 121 GTTGCCCTGTGCACCTAGTCTGGGAATTACACTACAGAGGCCACACAGTATGTTCTTGAT 180
|||||
QY 202 GlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLysAlaGluAspArgAla 221
|||||
Db 181 GGTGACAACCTCAGACATGGCTTTAATCGAGATCTAAGCTTCAAGGCAGACAGACCGTACA 240
|||||
QY 222 GluAsnIleArgAtqValGlyGluValAlaLysLeuPheAlaAspAlaGlyValIleCys 241
|||||
Db 241 GAAATATACGAGAGAGTTGGAGAGTGGCAAGCTTTTTCAGAGATGCTGGTACCATATGC 300
|||||
QY 242 IleAlaSerLeuIleSerProTyrArgArgAspArgAlaCysArgAlaLeuLeuPro 261
|||||
Db 301 ATTGCTAGTGTTCATATCTCCATACANGAGAGACCGGTGATGATGCCGAGCTCTACTTCCA 360
|||||
QY 262 HisSerAsnPhelIleGluValPheIleAspLeuProLeuLysIleCysGluAlaArgAsp 281
|||||
Db 361 GATTCTAGATTTATTGAAGTATTATGGAATTGCCACTAGAAATTATGTGAAGCTCGTGAT 420
|||||
RESULT 14
BF484142 746 bp mRNA linear EST 20-JUL-2000
LOCUS DGI_82_H05_g1_A002 Dark Grown 1 (DGI) Sorghum bicolor cDNA, mRNA
DEFINITION sequence.
ACCESSION BE361884
VERSION BE361884.1 GI:9303441
KEYWORDS EST.
SOURCE sorghum.
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE 1 (bases 1 to 746)
AUTHORS Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C., Sudman,M. and Pratt
,L.H.
TITLE An EST database from Sorghum: dark-grown seedlings
JOURNAL Unpublished (2000)
COMMENT Contact: Cordonnier-Pratt MM
Department of Botany
The University of Georgia
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 542 1805
Email: mmpratt@uga.edu
Sequences have been trimmed to exclude PolyA, vector and regions
below Phred quality 16. The threshold for highest quality sequence
is 20.
Seq primer: PolyTMix
High quality sequence start: 30
High quality sequence stop: 719
POLYA-No.
Location/Qualifiers
source 1..746

```

```

/organism="Sorghum bicolor"
/db_xref="taxon:4558"
/clone_lib="Dark Grown 1 (DGI)"
/notes="Organ: 5-day-old dark-grown seedlings; Vector:
Lambda Zap; Site.1: XhoI; Site.2: EcoRI; The library was
made from poly-A RNA in the cloning vector Lambda Zap II.
Clones to be sequenced were prepared by mass excision."
BASE COUNT 218 a 126 c 166 g 236 t
ORIGIN
Alignment Scores:
Pred. No.: 4,21e-31 Length: 746
Score: 604.00 Matches: 115
Percent Similarity: 96.72% Conservative: 3
Best Local Similarity: 94.26% Mismatches: 4
Query Match: 33.08% Indels: 0
DB: 10 Gaps: 0
US-09-720-384a-4 (1-343) x BE361884 (1-746)
QY 222 GluAsnIleArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyValIleCys 241
Db 2 GAAATATACGAAGAGTTGGTGAAGTGGCAAGCTTTTTCGCCGATCGTATCATATGC 61
QY 242 IleAlaSerLeuIleSerProTyrArgArgAspArgAspAlaCysArgAlaLeuPro 261
Db 62 ATGTGATCTGTGATATCTCCATACAGAGAGATCGTGATCGCGCTGCTACTTCCA 121
QY 262 HisSerAsnPheIleGluValPheIleAspLeuProLeuLysIleCysGluAlaArgAsp 281
Db 122 GATTCCTAATCTTATTGAAGTATTATGATTTGGCCCTAAATAATTTGTGAACCTGTGAT 181
QY 282 ProLysGlyLeuTyrLysLeuAlaArgThrGlyLysIleLysGlyPheThrGlyLeasp 301
Db 182 CTTAAGCTCTGTACAGCTTGCACGCACAGGAAGATTAAAGGTTTCACTGGAATTGAT 241
QY 302 AspProTyrGluProIleAsnGlyGluIleValIleLysMetLysAspGluLysCys 321
Db 242 GATCATACGAACGCCAGCTTAATGTCGAGATAGTAATTACGATGAAGATGGGAATGC 301
QY 322 ProSerProLysAlaMetAlaLysGlnValLeuCysTyrLeuGluAsnGlyTyrLeu 341
Db 302 CCTCACCCAAAGCAATGGCGAAGCAAGTCTCTATCATACCTTGAAGAGAACGGATATTG 361
QY 342 GlnAla 343
Db 362 CAAGCT 367
RESULT 15
BG852725
LOCUS 1024035X05.y2 C. reinhardtii CC-1690, normalized, Lambda Zap II
DEFINITION Chlamydomonas reinhardtii cDNA, mRNA sequence.
ACCESSION BG852725
VERSION 1
KEYWORDS EST.
SOURCE Chlamydomonas reinhardtii.
ORGANISM Chlamydomonas reinhardtii.
REFERENCE Grossman,A., Davies,J., Federspiel,N., Harris,E., Lefebvre,P.,
AUTHORS McDermott,J.P., Silflow,C., Stern,D. and Surzycki,R.
TITLE Analyses of the Chlamydomonas reinhardtii Genome: A Model,
Unicellular System for Analyzing Gene Function and Regulation in
Vascular Plants; project phase 2
JOURNAL Unpublished (2000)
COMMENT Contact: Charles Hauser
DCMB Box 91000
Duke University
Durham, NC 27708-1000
Tel: 919 613 8159
Fax: 919 613 8177

```

```

Email: chauser@duke.edu.
Location/Qualifiers
source 1..712
/organism="Chlamydomonas reinhardtii"
/strain="CC-1690 wild type mt+ 21gr"
/db_xref="taxon:3055"
/clone_lib="C. reinhardtii CC-1690, normalized, Lambda Zap
II"
/notes="Vector: pBluescript II SK-; Site.1: EcoRI; Site.2:
XhoI; This library, constructed by John Davies and Jeffrey
McDermott, combines cDNAs from CC-1690 cells grown to
mid-log phase in TAP (acetate-containing) medium in the
light, TAP medium in the dark, HS (minimal) medium in
ambient levels of CO2 and HS medium bubbled with 5% CO2.
PolyA mRNA was purified from each sample, pooled and cDNA
synthesized. The cDNA was directionally cloned into lambda
ZAP II (Stratagene) in the EcoRI (5') and XhoI (3') sites.
pBluescript II SK- plasmids were excised from the lambda
ZAP clones by superinfection with ExAssist (Stratagene)
phase. The library was normalized using method 4 described
in Ronaldo et al (1996) Genome Research 6: 791-806."
BASE COUNT 132 a 226 c 217 g 136 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 6,3e-31 Length: 712
Score: 601.00 Matches: 125
Percent Similarity: 67.36% Conservative: 36
Best Local Similarity: 52.30% Mismatches: 65
Query Match: 32.91% Indels: 14
DB: 10 Gaps: 3
US-09-720-384a-4 (1-343) x BG852725 (1-712)
QY 109 IleGlyArgTrpValArgArgArgArg-AsnGlyAlaAlaProGlyGluAlaProH 128
Db 17 ATCGCGCTTTGGCTCGCTCGCTCCGCCAATTTTCGCTGGCGCTCTGFGCCCTCAGC 76
QY 128 sSerProValLysGluLysProValMetSer-----AsnIleGlySerThrAs 145
Db 77 AGCT-----ATGAGCACTGCTATGACGTAGGATCTCTACCAA 115
QY 145 nIleLeuTrpHisAsnCysLeuIleGlyGlnSerAspArgGlnLysLeuGlyGlnLy 165
Db 116 CATTAAGTGGCAGCGGCGCTGTCCCGACGAAATAAGGAGCGGATTATGCAACAGAA 175
QY 165 sGlyCysValValTrpIleThrGlyLeuSerGlySerGlyLysSerThrLeuAlaCysAl 185
Db 176 GGGATCGCTGCTGTACCGGCTGACGGGCTGTGCAAGAGACACCGCTGCTGCAC 235
QY 185 aLeuSerArgGluLeuHisCysArgGlyHisLeuThrTyrValLeuAspGlyAspAsnLe 205
Db 236 GCTGGAGCAGCGCTCTTGGCGAGTCTGGCAAAAGTTACCGCTCTGCTGGACGCGACAACAT 295
QY 205 uArgHisGlyLeuAsnArgAspLeuSerPheLysAlaGluAspArgAlaGluAsnIleAr 225
Db 296 TCGGACGCGCTGAAACAGCAGCACTCGGATTCTCGCACCGATCGCGAGGAGCAATCCG 355
QY 225 gArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyValIleCysIleAlaSerLe 245
Db 356 CGCATTTGGCGAGGTCTCAAGGCTGTTCGCTGACACGGGTATTGCTACTGTGCTGAGCTT 415
QY 245 uIleSerProTyrArgArgAspArgAspAlaCysArgAlaLeuLeuProHisSerAsnPh 265
Db 416 CATCTCGCCCTACCGCGGAGCGGACCGCTGTGCGGCTCTCGCGTGGCGGAGGTGCTTT 475
QY 265 eIleGluValPheIleAspLeuProLeuLysIleCysGluAlaArgAspProLysGlyLe 285
Db 476 CATCGAGGTGTTATGAAGGTCTCCCATCTCTATCTGCGAGGAGCGGTGACCCCAAGGSCCT 535
QY 285 uTyrLysLeuAlaArgThrGlyLysIleLysGlyPheThrGlyIleAspAspProTyrGl 305
Db 536 CTACAAGAAGGCGCGCGCGGAGCTCAAAAGGCTTCCACCGCATTTGACGATCCCTACGA 595

```

Search completed: November 2, 2002, 03:05:16
Job time : 1693 secs